BJU Int 2022 doi:10.1111/bju.15740

Original Article

BJUI
BJU International

Development and validation of a life expectancy calculator for US patients with prostate cancer

Elizabeth C. Chase[1], Alex K. Bryant[2], Yilun Sun[3], William C. Jackson[2], Daniel E. Spratt[3], Robert T. Dess[2] and Matthew J. Schipper[1,2]

Department of [1]Biostatistics and [2]Radiation Oncology, University of Michigan, Ann Arbor, MI and [3]Department of Radiation Oncology, University Hospitals/Case Western Reserve University, Cleveland, OH, USA

R.T.D. and M.J.S. contributed equally as senior authors.

Objective

To develop and validate an accurate, usable prediction model for other-cause mortality (OCM) in patients with prostate cancer diagnosed in the United States.

Materials and Methods

Model training was performed using the National Health and Nutrition Examination Survey 1999–2010 including men aged >40 years with follow-up to the year 2014. The model was validated in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial prostate cancer cohort, which enrolled patients between 1993 and 2001 with follow-up to the year 2015. Time-dependent area under the curve (AUC) and calibration were assessed in the validation cohort. Analyses were performed to assess algorithmic bias.

Results

The 2420 patient training cohort had 459 deaths over a median follow-up of 8.8 years among survivors. The final model included eight predictors: age; education; marital status; diabetes; hypertension; stroke; body mass index; and smoking. It had an AUC of 0.75 at 10 years for predicting OCM in the validation cohort of 8220 patients. The final model significantly outperformed the Social Security Administration life tables and showed adequate predictive performance across race, educational attainment, and marital status subgroups. There is evidence of major variability in life expectancy that is not captured by age, with life expectancy predictions differing by 10 or more years among patients of the same age.

Conclusion

Using two national cohorts, we have developed and validated a simple and useful prediction model for OCM for patients with prostate cancer treated in the United States, which will allow for more personalized treatment in accordance with guidelines.

Keywords

calculator, comorbidities, life expectancy, other-cause mortality, prostate cancer, #PCSM, #prostate cancer, #uroonc

Introduction

Many men diagnosed with early-stage prostate cancer will not die from their cancer, instead dying from other diseases or natural conditions associated with advanced age [1]. Because competing risks of other-cause mortality (OCM) are substantial, they are a critical component of decision making for prostate cancer treatment. The National Comprehensive Cancer Network (NCCN) recommends that clinicians estimate patient other-cause life expectancy and incorporate these estimates into treatment decisions: in general, it is recommended that prostate cancer patients with life expectancy of 10 years or more receive the more aggressive treatment option appropriate to their cancer stage (e.g. surgery, radiation with hormonal therapy), while patients with a life expectancy of less than 10 years receive a less aggressive treatment option appropriate to their cancer stage (e.g. radiation alone, hormonal therapy alone, active surveillance) [2].

© 2022 The Authors.
BJU International published by John Wiley & Sons Ltd on behalf of BJU International. www.bjui.org wileyonlinelibrary.com
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Table 1 Literature review of life expectancy prediction tools in men with prostate cancer.

Reference	# Predictors (smaller better)	Web app? (preferred)	Provides continuous survival predictions (preferred)	Uses preformulated comorbidity index (not preferred)	Performance (bigger better)	Calibration	Notes
OCCAM (proposed)	8	Yes: occam-cap.org	Yes—both life expectancy and absolute risk predictions at requested timepoints between 0 and 15 years	No	External time-dependent AUC = 0.75 (10 years) External C-index = 0.70	External calibration shows slight pessimism (~1.5 years)	Data are from NHANES, 1999–2010. Used cause-specific Cox PH models for OCM endpoint. Validated in PLCO prostate cohort
Tan et al. *Journal of Urology*, 2021 [8]	34	No	NA	Yes—uses NCI comorbidity index as a predictor, along with others	Internal time-dependent AUC = 0.82 (10 years) No external validation	Calibration not discussed	Data are from SEER-CAHPS, 2004–2013. Used Fine & Grey models for OCM endpoint
Soerensen et al. *Urology*, 2021 [9]	90	No	NA	Yes—Care Assessment Needs score, which itself is built upon the Charlson Comorbidity Index	Internal time-dependent AUC = 0.74 (5 years) No external validation	Internal calibration at 5 years is satisfactory. No external calibration	Data are from Veterans Health Administration, 2013–2015. Used Cox PH models for OS endpoint
Sohlberg et al. *Urologic Oncology*, 2020 [10]	18	No	NA	Yes—builds model using the list of comorbidities from the Charlson Comorbidity Index	Internal C-index = 0.68 No external validation	Internal calibration is satisfactory. No external calibration	Data are from Veterans Health Administration, 2000–2015. Used Cox PH models for OS endpoint with predictors: age and number of comorbidities (from a list of 17 possible options)
Frendl et al. *PLOS One*, 2020 [11]	18	Yes: www.urologyrisk.com	NA	Yes—Charlson Comorbidity Index	Internal, bootstrapped C-index = 0.70 No external validation	Internal calibration is satisfactory. No external calibration	Data are from SEER-Medicare, 1998–2009. Used Fine & Grey models for OCM endpoint
Riviere et al. *JCO Clinical Cancer Informatics*, 2019 [12]	143	No	NA	No	External rho-squared = 0.68	Calibration not discussed	Data are from SEER-Medicare, 2004–2009. Used lasso for OS endpoint. Split data into training/testing sets to provide external validation
Daskivich et al. *Journal of Urology*, 2015 [13]	23	Explains how to calculate score by hand, if only using age + PCCI. (Validated performance is for age + PCCI + race, treatment, date of diagnosis, PSA, TNM stage, Gleason score)	No—just 2-, 5-, 10- and 15-year mortality	No	Partially external C-index = 0.77	Partially external calibration at 10 years is satisfactory	Training data are from Veterans Affairs hospitals, 1998–2004, and were used to develop PCCI score. Validation data are from Veterans Health Administration, 2000–2013. They fit Cox PH model *in validation data* with PCCI score, adjusted for age, race, treatment, date of diagnosis, PSA, TNM stage, Gleason score. Used OS endpoint
Kent et al. *BMC Medicine*, 2016 [14]	21	Yes: Memorial Sloan Kettering Male Life Expectancy Tool	No—just 10- and 15-year mortality	No	External C-index = 0.73	External calibration at 10 and 15 years shows slight pessimism	Used odds ratios calculated from the MALE model in a British private insurance database, then calibrated them using US Social Security Tables. Validated in PCOS cohort

Despite the recommendations on how to *incorporate* life expectancy into treatment decision making, little guidance on how to *estimate* other-cause life expectancy itself has been provided. The NCCN recommends that clinicians use Social Security Administration (SSA) actuarial tables to assess patient life expectancy. However, research suggests that the

© 2022 The Authors.

Table 1 (continued)

Reference	# Predictors (smaller better)	Web app? (preferred)	Provides continuous survival predictions (preferred)	Uses preformulated comorbidity index (not preferred)	Performance (bigger better)	Calibration	Notes
Hoffman et al. *J Gen Intern Med*, 2015 [15]	3	Nomogram presented	No—just 10- and 15-year mortality	No	Internal C-index = 0.73 No external validation	Calibration not discussed	Dataset was from 1994 to 1995, and surveyed men 6 months post diagnosis. Used Cox PH models for OS endpoint with predictors: age, race, and self-reported health status

Note that '# predictors' is the number of questions a clinician would have to ask the patient to use the tool. The blue fill in the top line draws attention to the fact that this is the proposed model. AUC, area under the curve; CAHPS, Consumer Assessment of Healthcare Providers and Systems; MALE, Measure of Actuarial Life Expectancy; NCI, National Cancer Institute; NHANES, National Health and Nutrition Examination Survey; OCCAM, other-cause comorbidity-adjusted mortality; OCM, other-cause mortality; OS, overall survival; PCCI, Prostate Cancer Comorbidity Index; PCOS, Prostate Cancer Outcomes Study; PH, proportional hazards; PLCO, Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial; SEER, Surveillance, Epidemiology and End Results.

SSA tables overestimate the life expectancy of patients with distant disease [3], and they do not adjust for patient comorbidities, which can have a notable effect on life expectancy [4]. To address these problems, the NCCN advised that clinicians combine SSA estimates with their clinical assessment of the patient's comorbidity burden. Unfortunately, research suggests that clinical intuition for estimating comorbidity burden is poor [3,5,6], and incorrect estimates of life expectancy may have ramifications on treatment recommendations, with high rates of overtreatment in patients with lower-risk prostate cancer and high comorbidity burden [7].

Existing tools to estimate comorbidity-adjusted life expectancy in men with prostate cancer have limitations. We identified eight life expectancy prediction tools in prostate cancer patients that used comorbidity information in some capacity to provide survival predictions (Table 1) [8–15]. All but one of these models requires clinicians to enter more than 10 pieces of information in order to use them, making them burdensome for routine use [8–14], and the remaining model (which uses three predictors) has not been externally validated [15]. Many of the models rely on preformulated comorbidity scores that were developed in populations with different characteristics from prostate cancer patients (e.g. women, young adults) or have not been updated to reflect outcomes of modern patients [8–11]. Only half are implemented as a web app or nomogram for clinical use [11,13–15]. No single tool had fewer than 10 predictors, was externally validated, and was implemented in a web app or nomogram. Therefore, we sought to develop and validate a simple, clinically usable prediction model for OCM in men with prostate cancer treated in the United States to personalize treatment decision making. We aimed to predict OCM and comorbidity-adjusted life expectancy accurately, and to improve on the SSA actuarial estimates and existing OCM prediction models.

Materials and Methods

Patients

Our training data originated from the National Health and Nutrition Examination Survey (NHANES), a cross-sectional survey of demographics and health conducted biennially in the United States since 1999. NHANES is nationally representative of the US non-institutionalized civilian population [16]. Mortality follow-up on NHANES participants is provided through linkage with the National Death Index, with follow-up through to 31 December 2014. NHANES is not limited to prostate cancer patients, but instead represents the entire US population. We used NHANES data from 1999 to 2010 and restricted these to men older than 40 years with complete data for all covariates used in model building: age; race; educational attainment; marital status; veteran status; insurance status; diagnosis of anaemia, angina, arthritis, asthma, chronic bronchitis, coronary heart disease, congestive heart failure, diabetes, emphysema, high cholesterol, hypertension, kidney issues, liver disease, mental health concern or myocardial infarction; alcohol use; healthcare access; whether the patient had been hospitalized in the past year; body mass index (BMI); and smoking status. Men who reported having ever had a malignancy diagnosis other than prostate cancer were excluded.

Our validation data originated from the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial (PLCO) prostate cohort. PLCO enrolled 155 000 participants in the United States between 1993 and 2001, with mortality follow-up to 31 December 2009. Participants were aged 55–74 years, with no history of prostate, lung, colon or ovarian cancer. Patients who met the trial inclusion criteria completed a medical history questionnaire at baseline and were randomized to receive cancer screening vs standard of care. If screening indicated a potential cancer, patients received diagnostic evaluation, which was included in the PLCO data along with

© 2022 The Authors.

primary treatment information [17]. For validation, we used the PLCO sample of men diagnosed with prostate cancer with complete data for the covariates in our candidate model.

Variable Definition

For model training, we used all-cause mortality (ACM) as our outcome, as few patients in the NHANES training population had prostate cancer, and the rate of prostate-cancer specific mortality (PCSM) in the general population is low (19.1 per 100 000 annually) [18]. Since PCSM comprises a small portion of ACM in men with localized prostate cancer, ACM is a reasonable approximation of OCM.

For model validation in PLCO, we used OCM as our primary endpoint to demonstrate the performance of our model to estimate OCM risk. To obtain OCM-specific performance estimates, men in the validation data who died from prostate cancer were censored at their time of death. This approach provides an estimate of cause-specific performance, which was our primary interest. However, we also considered performance in the subdistribution hazard setting, with PCSM treated as a competing risk event. More information, including details on variable definition, is provided in Appendix S1; the training data can be reconstructed using the code provided on GitHub.

Statistical Analysis

We considered three classes of models in model building: Cox proportional hazards models [19], survival random forests [20], and parametric cubic spline hazard models [21]. We used the NHANES survey weights when performing variable selection for the Cox models but did not use the survey weights in our final model or for model performance assessment. Including the survey weights complicated model performance assessment and was beyond the scope of this paper. All analyses were performed in R 3.6.2 [22].

Initial checks in the training data suggested that model performance would not suffer substantially from restricting to variables that also appeared in our validation data, so we focused on age, race, educational attainment, marital status, arthritis, chronic bronchitis, coronary heart disease, diabetes, emphysema, hypertension, BMI, and smoking status. We considered clinically relevant interactions and flexible functional forms for the effect of age. Using an internally cross-validated C-index to guide variable selection, we built separate models using each of our three modelling strategies in the NHANES data. Of these three, the Cox proportional hazards model exhibited the highest internally cross-validated C-index and time-dependent area under the curve (AUC) and was chosen as our final model to advance to testing in the external validation data. After model building but before advancing to external validation, we re-estimated the model effects and the baseline hazard in the training data using a larger sample of patients with complete data for the final predictors.

We conducted sensitivity analyses to assess whether our OCM prediction model developed in a non-prostate cancer population would translate to a prostate cancer patient population. In addition to considering prostate cancer diagnosis, years since diagnosis, and relevant interactions with these covariates as predictors in all models, we examined relationships between the linear predictor of our final model and prostate cancer diagnosis and years since diagnosis (all in NHANES data). These analyses suggested that having prostate cancer had little effect on predictions for ACM. Despite its minimal impact on model performance, we included prostate cancer as a predictor in our final model in order to adjust for its effect.

Using our candidate model and the baseline hazard estimated from NHANES, we obtained survival predictions for all men in PLCO and compared these predictions to their true outcomes. We calculated the inverse probability of censoring-weighted (IPCW) time-dependent AUC to assess discrimination and obtained confidence intervals for it using the variance estimator described by Blanche et al. (2013) [23]. Censoring weights were estimated using a marginal Kaplan–Meier model for censoring. We compared the IPCW time-dependent AUC of the SSA actuarial life tables, National Vital Statistics System (NVSS) life expectancy estimates, and our model using the multiple comparisons-adjusted testing procedure described by Blanche et al. (2013) [23], with timepoints of 5, 10 and 15 years. We also calculated the IPCW C-index; however, we used the time-dependent AUC as our primary metric of discrimination because of its superior properties in time-horizon settings [24]. We calculated the median survival (truncated at 15 years) as an estimate of life expectancy and made calibration plots to assess the calibration of these life expectancy predictions to observed patient life expectancy. We obtained OCM predictions from the 2001 SSA actuarial life tables and the NVSS life expectancy estimates. We compared the calibration plots of the SSA model to those of our final candidate model. The analysis and validation code are provided on GitHub; the validation data are not publicly accessible.

Algorithmic Bias

We performed analyses to assess the potential for algorithmic bias in our final model, guided by the framework outlined by Paulus and Kent [25]. We assessed the potential for algorithmic bias by estimating the IPCW time-dependent AUC stratified by race, educational attainment and marital status. Additionally, we estimated the proportion of patients who might have their life expectancy prediction appreciably changed by their marital status or educational attainment.

© 2022 The Authors.

Results

Cohort Features

After restricting the NHANES data to patients who met our inclusion criteria (Fig. S1), we were left with a training sample of 2420 men, of whom 459 died from all causes over a median of 8.8 years of follow-up among survivors. Of these deaths, 111 (24%) were from malignant neoplasms. Characteristics of the sample are given in Table 2. The mean age was 59 years. At the time of survey collection, 127 patients (5.2%) had been diagnosed with prostate cancer; of these, 18 had died from malignant neoplasms and 35 had died from other causes by the end of follow-up. Almost two-thirds of the sample were current or former smokers, and more than 75% of the sample had a BMI of ≥25 kg/m^2. The PLCO validation data consisted of 8220 patients with complete data for all predictors (Fig. S1). PLCO patients were markedly different from NHANES patients (Table 2). PLCO patients were older, more likely to be white, more educated, more likely to be married, and generally healthier than NHANES patients.

Table 2 Baseline characteristics of National Health and Nutrition Examination Survey (NHANES) training data and Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) validation data.

Predictor	Training (*N* = 2420)	Validation (*N* = 8220)
Mean (SD) age at diagnosis, years	59 (12)	70 (5.9)
Race, *n* (%)		
Non-Hispanic black	411 (17)	475 (5.8)
Non-Hispanic white	1362 (56)	7312 (89)
Other race	647 (27)	433 (5.3)
Education, *n* (%)		
Less than 9th grade	305 (13)	90 (1.1)
9th–11th grade	295 (12)	520 (6.3)
HS graduate	591 (24)	1512 (18)
Some college	607 (25)	2605 (32)
College graduate	622 (26)	3493 (43)
Marital status, *n* (%)		
Married	1813 (75)	7055 (86)
Separated	466 (19)	945 (12)
Single	141 (5.8)	220 (2.7)
Smoking status, *n* (%)		
Never	905 (37)	3368 (41)
Current	538 (22)	736 (9.0)
Former	977 (40)	4116 (50)
Arthritis: yes, *n* (%)	813 (34)	2471 (30)
Chronic bronchitis: yes, *n* (%)	111 (4.6)	240 (2.9)
Diabetes: yes, *n* (%)	453 (19)	527 (6.4)
Emphysema: yes, *n* (%)	75 (3.1)	183 (2.2)
Hypertension: yes, *n* (%)	1169 (48)	2742 (33)
Previous heart attack, coronary heart: yes, *n* (%)	281 (12)	979 (12)
Liver disease: yes, *n* (%)	113 (4.7)	314 (3.8)
Previous stroke: yes, *n* (%)	95 (3.9)	181 (2.2)
Body mass index, *n* (%)		
<18.5 kg/m^2	21 (0.9)	21 (0.3)
18.5–25 kg/m^2	506 (21)	2326 (28)
25–40 kg/m^2	1801 (74)	5809 (71)
40+ kg/m^2	92 (3.8)	64 (0.8)
Prostate cancer: yes, *n* (%)	127 (5.2)	8220 (100)
Outcome: deceased, *n* (%)	459 (19)	2415 (29)
Follow-up among survivors, median (range) months	105 (58, 180)	141 (9, 267)

Model Building and Validation

Our final other-cause comorbidity-adjusted mortality (OCCAM) model included eight predictors (age, diabetes, education, hypertension, marital status, smoking status, previous stroke and BMI) and interactions between age and diabetes, age and education, age and hypertension, and age and previous stroke. Model effect estimates for OCCAM are given in Fig. 1. Increased age, diabetes, hypertension, smoking, previous stroke, and having a BMI < 18.5 or greater than or equal 40 kg/m^2 all had harmful effects on OCM, while a higher education level and being married were protective. All model main effects were statistically significant. Age was the most important predictor, with time-dependent AUC at 10 years in the training set decreasing by 14% without age, followed by smoking status (0.7% decrease in AUC) and marital status (0.5% decrease). The least important predictor was prostate cancer (0.03% decrease). For transparency, we present the model based on all candidate predictors (pre-variable selection; no interaction terms) in Fig. S2 and the survey-weighted version of OCCAM in Fig. S3.

In external validation, OCCAM had a cause-specific IPCW AUC of 0.75 at 10 years and 0.78 at 15 years (Fig. S4). The IPCW C-index was 0.70. For comparison, the SSA life tables and NVSS life expectancy estimates produced IPCW AUCs of 0.71 and 0.75 at 10 and 15 years, respectively. The IPCW AUC of OCCAM was statistically significantly better than that of the SSA/NVSS at 5, 10 and 15 years ($P < 0.001$ at all timepoints). Estimates of the IPCW AUC for a subdistribution hazard interpretation are given in Table S1; performance was similar to performance in the cause-specific setting.

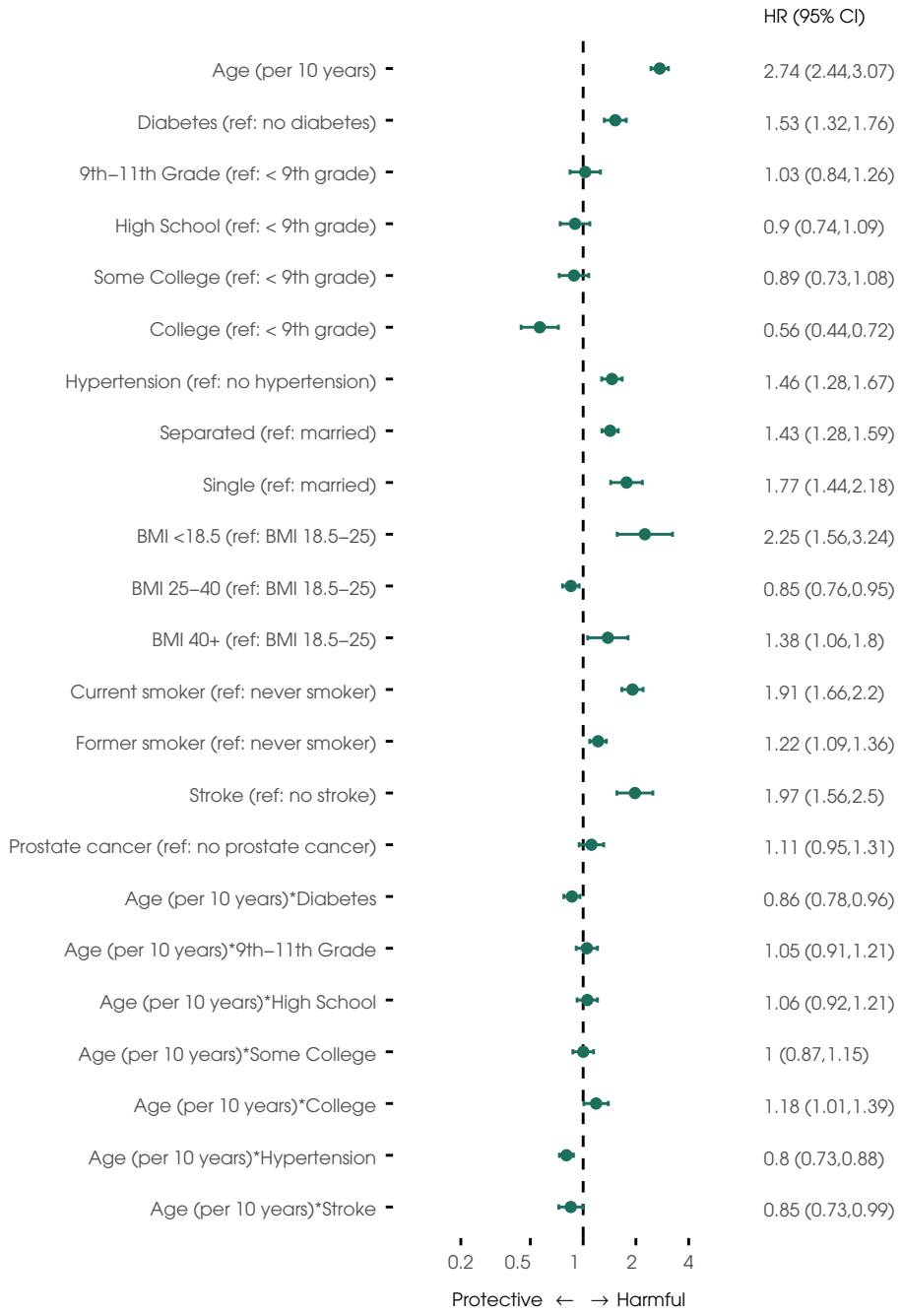
Calibration of OCCAM was substantially better than the SSA estimates, which showed pessimism of approximately 3–6 years (Fig. 2). OCCAM, by comparison, shows pessimism of approximately 1.5 years, which may be attributable in part to the extremely healthy nature of the PLCO population relative to NHANES. On the risk scale, the calibration of OCCAM was generally good at 15 years; performance was weaker at 10 years, with worst performance for patients with predicted OCM of 70% at 10 years, for whom there was pessimism of approximately 20% (Fig. S5).

Algorithmic Bias

Cause-specific IPCW AUC, stratified by race, marital status and educational attainment, is given in Tables S2–S4. Model discrimination was reduced for non-Hispanic Black men relative to non-Hispanic White men and men of other race at

© 2022 The Authors.

Fig. 1 Forest plot of predictors for other-cause comorbidity-adjusted mortality (OCCAM) model fit in the National Health and Nutrition Examination Survey training cohort of 7369 men. HR, hazard ratio.



5 and 10 years; all groups had similar predictive performance by 15 years. When stratified by marital status, OCCAM exhibited improved predictive performance for separated men (those divorced or widowed) at 5 years relative to men who were married/in a live-in relationship or single; performance was similar by 10 and 15 years. When stratified by educational attainment, men with a 9th–11th grade education had slightly reduced predictive performance at all time points relative to other groups. Men with lower than a 9th grade education level had the best predictive performance at 5 years and the worst predictive performance at 10 and 15 years; men with some college or a college degree had the best predictive performance at 10 and 15 years.

We found that changing marital status to the most protective option (being married or currently living with a partner)

© 2022 The Authors.

Fig. 2 Calibration performance of other-cause comorbidity-adjusted mortality model (OCCAM) and the Social Security Administration's 2001 actuarial life table predictions (SSA) in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial cohort of 8220 men with prostate cancer.

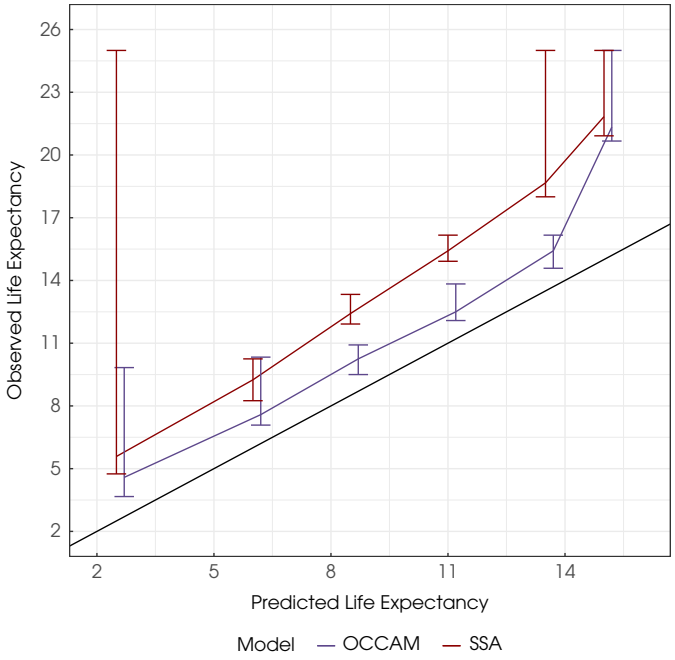


Fig. 3 Scatterplot comparing Social Security Administration's 2001 actuarial life table predictions (SSA) and patient age to predictions from our other-cause comorbidity-adjusted model (OCCAM) in a cohort of 8220 men with prostate cancer from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO). Concordance is defined here as whether or not it would alter treatment according to National Comprehensive Cancer Network guidelines, which makes treatment determinations based on whether or not life expectancy is greater or less than 10 years, marked here by horizontal and vertical lines. The diagonal line indicates perfect agreement.

altered life expectancy prediction by less than 1 month for almost 90% of patients in NHANES and PLCO, and fewer than 3% of patients had their life expectancy changed from <10 years to ≥10 years. Changing educational attainment to the most protective option (college graduate) altered life expectancy prediction by less than 1 month for 65% of patients in NHANES and PLCO, and fewer than 5% of patients would switch from the <10 years to ≥10 years group.

Other-cause Mortality Risk Drivers

There was substantial variation in life expectancy within age (Fig. 3). For a 74-year-old man, SSA predicted life expectancy was 10 years, but OCCAM predictions ranged from 5 years to more than 15 years. For younger men, SSA predictions were more optimistic because they failed to account for comorbidities that reduce life expectancy. For older men, SSA predictions were pessimistic because they failed to recognize that elderly men can be very healthy. Based on the NCCN threshold of 10-year life expectancy for particular treatment recommendations, these discrepancies between OCCAM and SSA altered NCCN treatment recommendations in approximately 15% of the men in PLCO. Twelve percent of men in PLCO would have their NCCN-recommended treatment altered by an over-pessimistic prediction from SSA, while 2.4% of men in PLCO would have their treatment altered by an over-optimistic prediction from SSA.

Patient characteristics, split by OCCAM-predicted life expectancy, can be found in Table S5. Patients with reduced other-cause life expectancy were more likely to be Black or of other race. Graduating from college was highly protective for OCM, as was being married. Current smokers and former smokers were overrepresented in the high OCM risk group, as were patients with BMI < 18.5 and ≥40 kg/m^2. Prostate cancer characteristics were also linked with OCCAM predictions. Patients with more aggressive cancers were overrepresented in the high OCM risk group, while patients with less aggressive cancers generally had reduced OCM risk. Sample sizes for aggressive cancers were quite low, given the nature of the PLCO trial, therefore, conclusions about these patients must be made with care. It is also important to note the overtreatment in PLCO, with many patients treated aggressively even for early-stage cancer. OCCAM predictions can be further explored using the app given at occam-cap.org.

Discussion

Using two prospective US national cohorts, we developed and validated OCCAM, a prediction model for OCM in prostate

© 2022 The Authors.

cancer patients. Our model performs well despite being validated in a population that was quite different from its training population. Furthermore, it only requires eight predictors. OCCAM adds to the list of OCM prediction models for men with prostate cancer; however, it is unique in its high externally validated predictive performance, simplicity, and usability in diverse patient populations [8–15]. OCCAM estimates other-cause life expectancy, which allows clinicians to follow NCCN guidelines, and because of its simplicity and accessibility, it facilitates future research on OCM in prostate cancer patients.

When building OCCAM, many predictors and modelling strategies were considered. The statistical methods used to produce the model are rigorous and do not rely on intermediate risk scores. We present both life expectancy and risk predictions over time, rather than point estimates. These features make it attractive for future statistical work, and because we provide all training data and code used to produce the model, OCCAM can be combined with other prediction tools and treatment effect estimates to build new models. The approach we use here—building an OCM model for cancer patients in a non-cancer patient study—may be generalizable to other cancer patient populations where the cancer-specific mortality is low.

We provide evidence that incorporating comorbidity information is beneficial to predicting OCM in prostate cancer patients. OCCAM significantly outperformed the SSA life tables and NVSS life expectancy estimates. In terms of absolute performance, age-only predictors do perform well and provide a benchmark for future models. For research purposes, two of our predictors—educational attainment and marital status—pose problems, as they are not routinely collected in prostate cancer cohorts. In post hoc checks, we found that our model refit in NHANES without education and marital status still performed quite well in external validation, with a time-dependent AUC of 0.74 at 10 years (Fig. S6).

Educational attainment and marital status pose an additional problem: the potential for algorithmic bias [25]. Although both predictors may have a direct predictive effect on OCM, they are also likely predictive in part because of their association with race and racism, socioeconomic status, and health insurance access. One approach to address this issue is to consider a model without education and marital status, which as described, had good predictive performance in post hoc checks. However, changing our model post validation may not sufficiently address the problem. As Paulus and Kent note, excluding problematic predictors is not a cure-all and could worsen injustice. The resulting model averages over the problematic predictor, and thus predictions are implicitly predictions for the majority [25]. Our ultimate goal is to provide accurate predictions of other-cause life expectancy for all prostate cancer patients, to serve as one factor in the complex decision on which treatment course is appropriate for their cancer and any other medical problems they have. Giving the same prediction for all men regardless of their educational attainment and marital status risks giving some men over-optimistic or over-pessimistic predictions that could unnecessarily subject them to aggressive treatment or fail to give them life-extending care.

To assess algorithmic bias, we followed the framework outlined by Paulus and Kent for advancing algorithmic justice in healthcare prediction models [25]. First, we found that, although there were differences in predictive performance across subgroups, predictive performance was similar across all subgroups. The reduced predictive performance for non-Hispanic Black men at 10 years (AUC = 0.67) is the largest cause for concern. That may be attributable to the small number and unique features of Black men in the PLCO external validation data (n = 475, or 5.8% of the sample), which reinforces the need for broad diversity and inclusion in future trials. Second, we found that fewer than 5% of patients in NHANES or PLCO would have had their prediction appreciably changed by a modification to their marital status or educational attainment. Given the diverse range of patients included in NHANES and PLCO, these findings suggest that marital status and educational attainment are not the primary factor behind a treatment decision for most patients. Third, to ensure that users are aware when marital status or educational attainment affect treatment decisions, our app gives the option to generate predictions using either the original OCCAM, or the model re-estimated without either marital status, educational attainment, or both. For all patients, it calculates whether a change in marital status or educational attainment would appreciably change a patient's life expectancy prediction, and, if so, prints a warning for the app user and recommends looking at the prediction for when the 'Do not consider marital status/educational attainment' option is checked. Including these predictors in the app—and the option to exclude them—may generate conversations between patients and healthcare providers about barriers to care caused by marital status and educational attainment (e.g. treatment expense, uncertainty or misunderstanding about a particular treatment, lack of caregiving support) that could assist clinicians in providing holistic care.

Finally, our model is still subject to what Paulus and Kent call labelling bias, in which the outcome is inherently different across groups [25]. In our case, we may implicitly be modelling 'life expectancy under optimal care' for some patients (perhaps wealthier, white, or more educated patients) while modelling 'life expectancy under suboptimal care' for other patients (perhaps less affluent, less educated, or Black or Indigenous people of colour), because of disparities in access to high-quality care. Labelling bias cannot be identified or ruled out from the data and is a product of systemic

© 2022 The Authors.

inequity in our society. We cannot correct for it, but readers should be aware of this issue [25].

Beyond the potential for algorithmic bias, which we have done our best to mitigate, our study is subject to several limitations. Building a cancer OCM prediction model in a non-cancer patient population could introduce bias. We addressed this limitation through sensitivity analyses of the effect of prostate cancer within our model, and we validated our model in a prostate cancer population. Some may argue that our use of cause-specific hazard modelling instead of Fine and Grey subdistribution hazard modelling is a weakness, necessitated by NHANES not being a prostate cancer patient population. However, we argue that the cause-specific hazard approach is more reflective of clinical use, and we wanted a cause-specific model in order to provide integrated cumulative incidence predictions in future work [26,27]. There may be some concerns about NHANES' reliance on patient self-report. We believe that patient self-report may be more reflective of clinical use, because physicians may choose to ask patients directly, rather than relying on the patient's medical record. More work is needed to optimally deploy OCCAM for routine clinical use. We have not studied patient perceptions of the OCCAM app or studied how it functions in clinical settings. We hope that future research can give more insight into OCCAM's performance in actual use.

Our other-cause comorbidity-adjusted mortality model exhibited some pessimism (approximately 1.5 years) when validated in PLCO. Although an improvement over the SSA tables (which exhibited pessimism of 3–6 years) and comparable to calibration of other models that have undertaken external calibration validation [14], this pessimism is still important to note. We suspect that this may in part be caused by our use of PLCO validation data, as the PLCO population is an unusually affluent, healthy prostate cancer patient population that is not representative of the general prostate cancer population. PLCO exhibits high rates of overtreatment and low rates of PCSM compared to other prostate cancer patient populations, and thus it is hard to assess the true calibration of our model for all prostate cancer patients based on PLCO alone. In addition, Pierre-Victor et al. [28] found that there was a relationship between PCSM and OCM in PLCO, with patients with higher-risk disease more likely to experience OCM than patients with lower-risk disease. We saw evidence of this relationship in our findings (Table S2). This association could cause bias when trying to separate these two endpoints.

Despite these drawbacks of PLCO, there are very few prostate cancer patient datasets that contain PLCO's richness of prostate cancer *and* other-cause comorbidity predictors. We hope that this and related work encourage the collection of more representative prostate cancer patient data that examine both PCSM and OCM risk factors, and we look forward to validating OCCAM in more datasets in future. Until OCCAM is validated in more generalizable datasets, its external validity outside of NHANES and PLCO is uncertain.

We have built and validated OCCAM, an accurate and succinct model for OCM in US prostate cancer patients. Our model requires only eight inputs and fills a clinical void: simple, accurate models to predict OCM are lacking, despite the fact that national guidelines recommend incorporating patient life expectancy into treatment decisions for men with localized prostate cancer. OCCAM can be used in accordance with NCCN guidelines and has high potential to improve quality of care when patient life expectancy is a factor.

Acknowledgements

We thank the Cancer Data Access System staff for their assistance in obtaining our validation data. This work was supported by the National Science Foundation (grant number DGE-1256260) and the National Cancer Institute at the National Institutes of Health (grant number P30-CA046592).

Disclosure of Interests

Dr Schipper consults for Innovative Analytics (unrelated to submitted work). Dr Spratt receives personal fees from Bayer, Boston Scientific, Janssen, GammaTile, Blue Earth, Pfizer, Myovant, Novartis and Varian. All other authors have no conflicts of interest to report.

References

1 **American Cancer Society. Key Statistics for Prostate Cancer**. https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html. Updated 2019. Accessed September 2020
2 **National Comprehensive Cancer Network**. NCCN Guidelines Version 2.2020: prostate cancer. Plymouth Meeting, PA: National Comprehensive Cancer Network; 2020
3 **Sammon JD, Abdollah F, D'Amico A et al.** Predicting life expectancy in men diagnosed with prostate cancer. *Eur Urol* 2015; 68: 756–65
4 **Jeldres C**. Life expectancy estimation in prostate cancer patients. *Can Urol Assoc J* 2012; 6: 374–5
5 **Leung KMYB, Hopman WM, Kawakami J**. Challenging the 10-year rule: the accuracy of patient life expectancy predictions by physicians in relation to prostate cancer management. *Can Urol Assoc J* 2012; 6: 367–73
6 **Walz J, Gallina A, Perrotte P et al.** Clinicians are poor raters of life-expectancy before radical prostatectomy or definitive radiotherapy for localized prostate cancer. *BJU Int* 2007; 100: 1254–8
7 **Daskivich TJ, Chamie K, Kwan L et al.** Overtreatment of men with low-risk prostate cancer and significant comorbidity. *Cancer* 2011; 117: 2058–66
8 **Tan HH, Zhou X, Spratte BN et al.** Patient reported vs claims based measures of health for modeling life expectancy in men with prostate cancer. *J Urol* 2021; 205: 434–40
9 **Soerensen SJC, Thomas IC, Schmidt B et al.** Using an automated electronic health record score to estimate life expectancy in men diagnosed with prostate cancer in the veterans health administration. *Urology* 2021; 206: e211
10 **Sohlberg EM, Thomas IC, Yang J et al.** Life expectancy estimates for patients diagnosed with prostate cancer in the veterans health administration. *Urol Oncol* 2020; 38: 734.e1–10

© 2022 The Authors.

11 **Frendl DM**, **FitzGerald G**, **Epstein MM et al.** Predicting the 10-year risk of death from other causes in men with localized prostate cancer using patient-reported factors: development of a tool. *PloS One* 2020; 15: e0240039

12 **Riviere P**, **Tokeshi C**, **Hou J et al.** Claims-based approach to predict cause-specific survival in men with prostate cancer. *JCO Clin Cancer Inform* 2019; 3: 1–7

13 **Daskivich TJ**, **Kwan L**, **Dash A et al.** An age adjusted comorbidity index to predict long-term, other cause mortality in men with prostate cancer. *J Urol* 2015; 194: 73–8

14 **Kent M**, **Penson DF**, **Albertsen PC et al.** Successful external validation of a model to predict other cause mortality in localized prostate cancer. *BMC Med* 2016; 14. https://doi.org/10.1186/s12916-016-0572-z

15 **Hoffman RM**, **Koyama T**, **Albertsen PC et al.** Self-reported health status predicts other-cause mortality in men with localized prostate cancer: results from the prostate cancer outcomes study. *J Gen Intern Med* 2015; 30: 924–34

16 **National Center for Health Statistics**, **Centers for Disease Control and Prevention**. National Health and Nutrition Examination Survey Data, 1999–2010. https://wwwn.cdc.gov/nchs/nhanes/ContinuousNhanes/Default.aspx. Updated 2019. Accessed June, 2019

17 **Division of Cancer Prevention**, National Cancer Institute. Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO). https://prevention.cancer.gov/major-programs/prostate-lung-colorectal-and-ovarian-cancer-screening-trial. Updated 2019. Accessed August, 2019

18 **U.S. Cancer Statistics Working Group**, **Centers for Disease Control and Prevention and National Cancer Institute**, **U.S. Department of Health and Human Services**. U.S. Cancer Statistics Data Visualizations Tool, based on 2019 submission data (1999–2017). www.cdc.gov/cancer/dataviz. Released June 2020. Accessed September, 2020

19 **Cox DR**. Regression models and life tables. *J R Stat Soc Series B Stat Methodol* 1972; 34: 187–220

20 **Ishwaran H**, **Kogalur UB**, **Blackstone EH et al.** Random survival forests. *Ann Appl Stat* 2008; 2: 841–60

21 **Royston R**, **Parmar MKB**. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002; 21: 2175–97

22 **R Core Team**. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2019. https://www.R-project.org/. Accessed June, 2019

23 **Blanche P**, **Dartigues JF**, **Jacqmin-Gadda H**. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013; 12: 5381–97

24 **Blanche P**, **Kattan MW**, **Gerds TA**. The c-index is not proper for the evaluation of t-year predicted risks. *Biostatistics* 2019; 20: 347–57

25 **Paulus JK**, **Kent DM**. Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *npj Digit Med* 2020; 3: 99

26 **Prentice RL**, **Kalbfleisch JD**, **Peterson AV et al.** The analysis of failure times in the presence of competing risks. *Biometrics* 1978; 34: 541–54

27 **Pfeiffer RM**, **Gail MH**. *Absolute risk: methods and applications in clinical management and public health*. London: Chapman & Hall/CRC Monographs on Statistics and Applied Probability, 2017

28 **Pierre-Victor D**, **Pinsky PF**, **Miller E et al.** Other and all-cause mortality among men diagnosed with prostate cancer in the PLCO trial. *J Urol* 2021; 205: 1372–8

Correspondence: Elizabeth C. Chase, Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA.

e-mail: ecchase@umich.edu

Abbreviations: ACM, all-cause mortality; AUC, are under the curve; BMI, body mass index; IPCW, inverse probability of censoring-weighted; NCCN, National Comprehensive Cancer Network; NHANES, National Health and Nutrition Examination Survey; NVSS, National Vital Statistics System; OCCAM, other-cause comorbidity-adjusted mortality; OCM, other-cause mortality; PCSM, prostate-cancer specific mortality; PLCO, Prostate, Lung, Colon, and Ovarian Cancer Screening Trial; SSA, Social Security Administration.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Additional information on data sources, model building and model validation.

Fig. S1. STROBE diagram to construct National Health and Nutrition Examination Survey training data and Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial validation data.

Fig. S2. Forest plot from full model using all candidate predictors, both unweighted and survey weighted, fit in the National Health and Nutrition Examination Survey training cohort of 2420 men.

Fig. S3. Forest plot comparing the effect sizes of predictors from the unweighted and survey weighted versions of our other-cause comorbidity-adjusted mortality model (OCCAM) fit in the National Health and Nutrition Examination Survey (NHANES) training cohort of 7369 men.

Fig. S4. Externally validated time-dependent areas under the curve at 5, 10 and 15 years of our other-cause comorbidity-adjusted mortality model, the Social Security Administration 2001 actuarial life table predictions, and the National Vital Statistics System's 2001 life expectancy predictions. Models were validated in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial cohort of 8220 men with prostate cancer.

Fig. S5. Calibration performance of other-cause comorbidity-adjusted mortality model (OCCAM) for 10- and 15-year risk of other-cause mortality (OCM) in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial cohort of 8220 men with prostate cancer.

Fig. S6. Forest plot of the reduced other-cause comorbidity-adjusted mortality (OCCAM) model without education and marital status predictors, for both the unweighted and survey weighted versions of the model fit in the National Health and Nutrition Examination Survey training cohort of 7369 men.

© 2022 The Authors.

Table S1. Cause-specific and competing-risks time-dependent areas under the curve (AUCs) of the other-cause comorbidity-adjusted model (OCCAM) in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial validation cohort of 8220 men.

Table S2. Cause-specific time-dependent areas under the curve (AUCs) of the other-cause comorbidity-adjusted mortality (OCCAM) model in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial validation cohort of 8220 men, stratified by race.

Table S3. Cause-specific time-dependent areas under the curve (AUCs) of the other-cause comorbidity-adjusted mortality (OCCAM) model in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial validation cohort of 8220 men, stratified by marital status.

Table S4. Cause-specific time-dependent areas under the curve (AUCs) of the other-cause comorbidity-adjusted mortality (OCCAM) model in the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial validation cohort of 8220 men, stratified by educational attainment.

Table S5. Characteristics of a cohort of 7596 men from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial, grouped by our other-cause comorbidity-adjusted mortality model (OCCAM)-predicted median survival time.

© 2022 The Authors.