Development and Validation of a Clinically Usable Prediction Model for Other-Cause Mortality in Men with Prostate Cancer using Two Prospective National Cohorts

**Abstract**

**Background**:

**Methods**:

**Findings**:

**Interpretation**:

**Funding**: National Science Foundation (DGE-1256260)

Word Count

Abstract (max 275):

Main text (max 3000): 3411

6 figures + tables (currently at 6-7… ish?)

**Introduction**

In 2019, there were 174,650 new cases of prostate cancer in the US. However, most men diagnosed with prostate cancer will not die of their cancer, instead dying of old age or other comorbidities [1]. Because competing risks of other-cause (OC) mortality are substantial, they may render prostate cancer treatment or treatment intensification futile. As a result, prostate cancer treatment guidelines recommend that clinicians take the patient's risk of OC mortality and overall life expectancy into account when making treatment decisions.

Little guidance on how to estimate other-cause mortality in prostate cancer patients has been provided. The National Comprehensive Cancer Network (NCCN) recommends that clinicians use Social Security Administration (SSA) actuarial tables to assess patient life expectancy [2]. The SSA tables have several weaknesses. First, research has shown that prostate cancer patients with localized disease usually live 4-6 years longer than SSA predictions would suggest, while the SSA predictions overestimate the life expectancy of patients with distant disease [3]. Second, the SSA tables do not adjust for patient comorbidities, although the life expectancy of a 70-year-old man may vary by more than 11 years from his SSA predicted expectancy depending on his comorbidity burden [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]. In two studies that looked at physician performance for estimating prostate cancer patient life expectancy, both found that physicians correctly predicted whether a patient would be alive in ten years only 68% of the time [5, 6]. It should go without saying that some physicians were better than others, with predictive performance across individual physicians ranging from 52% to 78%, another weakness of relying on physician opinion [6]. These incorrect estimates of patient life expectancy may have major ramifications on treatment plan: Daskivich et al. report a high rate of overtreatment in patients with low-risk prostate cancer and high comorbidity burden, which may have negative effects on patient overall mortality and quality of life [7].

In response to these challenges, researchers have worked to build and validate OC mortality prediction tools in men with prostate cancer. Based on a literature review, we identified four OC mortality prediction models specifically for men with prostate cancer that provided survival predictions and incorporated comorbidity information in some capacity [8, 9, 10, 11]. However, we had concerns about all four models. Hoffman et al. (2015) built a model with C-index of 0.73 using demographic information, treatment information, Gleason score, and self-reported health status [8]. We questioned if model performance could be further improved by incorporating more comorbidity information beyond self-reported health status, and the inclusion of treatment information as a predictor makes it difficult to use the tool to advise treatment decisions. Daskivich et al. (2015) built a model that age-adjusted the Charlson Comorbidity Index (CCI) and calibrated it to men with prostate cancer, obtaining a C-index of 0.77 in the veteran population [9]. We wanted to incorporate comorbidities without the constraint of a pre-formulated comorbidity score.

Kent et al. (2016) combined a comorbidity-life expectancy tool built in a British private insurance claims database with age-shifted SSA tables; their model had a C-index of 0.73 [10]. We had concerns about some of the statistical choices Kent et al. made and whether a British private insurance population is representative of the general US patient population. In addition, the model requires 19 specialized predictors, many of which may not be known by the average patient, or even appear in his medical record, thus necessitating additional diagnostic workup in order to use the model. Finally, Riviere et al. (2019) built an OC mortality prediction model in SEER-Medicare claims data, reporting a of 0.6. Their final model involved 143 covariates, which we thought would be burdensome in a clinical setting [11]. We were also concerned by their use of SEER-Medicare claims data, which may produce models that do not generalize well to non-claims settings. For example, health conditions like obesity and smoking are often underdiagnosed in claims data, because they do not generate an insurance claim unless very severe. One study reports the prevalence of obesity in SEER-Medicare cancer patients as 0.7%, clearly not true of the general U.S. population [12]. This can lead to unexpected biases and generalizability issues.

We seek here to develop a clinically usable, validated prediction model for OC mortality in US men with prostate cancer to help personalize treatment decision making. We use twelve years of data from the National Health and Nutrition Examination Survey (NHANES) to build our prediction model and validate it in the prostate cohort of the Prostate, Lung, Colon, and Ovarian Cancer Screening Trial (PLCO). Our final model is parsimonious, does not rely on treatment information, and is built and validated in a diverse sample that reflects most U.S. patients with localized disease assessed in a clinical setting. We have made it available for public use in app format at https://elizabethchase.shinyapps.io/pctreat/.

**Methods**

*Data Acquisition*

Our training data come from NHANES, a cross-sectional survey conducted every two years in the US since 1999. NHANES is nationally representative of the U.S. non-institutionalized civilian population, and contains information on patient demographics and health. More information on NHANES' sample construction and data collection protocols can be found at the survey's website [13]. Mortality follow-up on all NHANES participants is provided through a linkage with the National Death Index, with mortality follow-up through Dec. 31, 2014. Critically, NHANES is not limited to prostate cancer patients, but instead represents the entire US population.

We used NHANES data from 1999-2010. We restricted to men older than 40 free of non-prostate cancer with complete data for all covariates used in model-building: age, race, educational attainment, marital status, veteran status, insurance status; diagnosis of anemia, 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; and alcohol use, regular access to healthcare, whether the patient had been hospitalized in the past year, BMI, and smoking status. More information on variable definition is provided in the supplementary materials; the training data can be reconstructed using the code provided at https://github.com/elizabethchase/PCOtherCause.

*Model Building*

We considered three classes of models in initial model building: Cox proportional hazards models [14], survival random forests [15], and parametric cubic spline hazard models [16]. In all cases we used overall mortality as our outcome (as opposed to OC-specific mortality) because most patients in our population did not have prostate cancer and thus could not be censored for dying of it. This assumption is justifiable because prostate-cancer specific mortality comprises only a small portion of overall mortality in men with prostate cancer. We did sensitivity analyses, explained further below, to further support this choice. We used the NHANES survey weights when performing variable selection for the Cox models, but otherwise did not weight our analysis. Including the survey weights complicated model performance assessment, and statistical methodology to incorporate survey weights into survival random forest and parametric spline modeling is beyond the scope of this paper. All analyses were performed in R 3.6.2 [17].

Initial checks suggested that model performance would not suffer substantially from restricting to variables that also appeared in PLCO, so we focused on age, race, educational attainment, marital status, arthritis, chronic bronchitis, coronary heart disease, diabetes, emphysema, hypertension, BMI, and smoking status. Using a combination of clinical knowledge, statistical significance, and assessment of cross-validated C-index, we built separate models using each of our three modeling strategies in the NHANES data. We also considered models built in the age 55+ population (to match the validation data's inclusion criteria) and age-weighted to match the PLCO age structure.

We conducted sensitivity analyses to gauge the effect of estimating a prostate cancer OC mortality prediction model in a non-prostate cancer 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 models and prostate cancer diagnosis and years since diagnosis. These analyses all suggested that having prostate cancer had little effect on predictions for overall mortality. Despite its minimal impact on model performance, we still included prostate cancer as a predictor in all models in order to adjust for its effect.

After completing sensitivity analyses and model-building, we used the internally cross-validated C-index to settle on three candidate models to advance to validation in PLCO: a Cox proportional hazards model fit in men older than 40, a second Cox proportional hazards model fit in men ages 55+, and a survival random forest fit in men older than 40. Model effects, baseline hazards, and survival trees were re-estimated in the training data using a larger sample of patients with complete data for the final group of predictors before advancing to validation. Modeling code is provided at https://github.com/elizabethchase/PCOtherCause.

*Model Validation*

We validate our model in the PLCO prostate cohort. PLCO enrolled 155,000 participants between November 1993 and July 2001, with mortality follow-up through Dec. 31, 2009. In order to enroll in the trial, participants had to be ages 55-74, with no history of prostate, lung, colon, or ovarian cancer, and with prostate, lungs, colon, and ovaries intact. Additionally, participants could not be receiving treatment for any cancer, and they could not have received screening for prostate, lung, colon, or ovarian cancer in the past year. Patients who met trial inclusion criteria completed a medical history questionnaire at baseline and were randomized to receive screenings for prostate, lung, and colon cancers if male. If a screening indicated a potential cancer, patients received diagnostic workup, which was included in the PLCO data along with their primary treatment information. For more information on the PLCO protocol and data, please see the trial's website [18].

For validation, we restricted the PLCO sample to men who had been diagnosed with prostate cancer with complete data for the covariates in our three candidate models. We censored men who died of prostate cancer. Using the candidate models, we obtained survival predictions for all men in PLCO and compared these predictions to their true mortality outcomes. We calculated the C-index and time-dependent AUC and selected the best performing model as our final OC mortality prediction model. Using calibration plots, we assessed the calibration of our cause-specific model built in NHANES and a Fine and Gray regression fit in the PLCO data using the NHANES linear predictor as the sole predictor.

Using the Fine and Gray model, we grouped the PLCO sample into OC mortality risk groups and present descriptive statistics on these patients. We use these OC mortality predictions in combination with information on patients’ prostate cancer characteristics to make treatment recommendations consistent with NCCN guidelines and compare these recommendations with the treatments that patients actually received. Validation code is provided on GitHub; the validation data are not publicly accessible and cannot be provided.

**Results**

*Model Building and Assessment*

After restricting the NHANES data to patients who met our inclusion criteria (Figure 1), we were left with a training sample of 2,420 men, of whom 459 died over a mean of 103.7 months (8.6 years) of follow-up. Characteristics of the sample are given in Table 1. Mean age was 59.4 years. 127 patients (5.2%) had been diagnosed with prostate cancer at the time of survey collection. Almost two-thirds of the sample were current or former smokers, and more than 75% of the sample was overweight or obese. The PLCO external validation data consisted of 8,220 patients with complete data for all predictors (Figure 2). PLCO patients were markedly different from NHANES patients (Table 1). PLCO patients were older, more likely to be white, more educated, more likely to be married, and generally healthier than NHANES patients. The only characteristic on which NHANES and PLCO patients were not significantly different was previous heart attack or diagnosis of coronary heart disease; roughly 12% of patients had this diagnosis in both samples.

Our age 40+ Cox model included eight predictors (age, diabetes, education, hypertension, marital status, smoking status, previous stroke, and BMI); our age 55+ Cox model included nine (age, race, education, marital status, emphysema, diabetes, previous stroke, smoking status, and BMI); and our random forest included fourteen (age, arthritis, chronic bronchitis, diabetes, education, emphysema, hypertension, marital status, previous heart attack, BMI, liver disease, race, smoking status, and previous stroke). All three models were refit using all patients with complete data for this shorter list of covariates; for the age 40+ Cox model, this was a sample of 7,369 men; for the age 55+ Cox model, a sample of \_\_\_\_ men; for the random forest, a sample of \_\_\_\_ men. Descriptive statistics on these three enlarged training samples are given in the Supplementary Materials.

The age 40+ Cox model performed best in external validation (Table 2), with a C-index of 0.70 and an AUC of 0.78 at 14 years. Model performance was largely similar across prostate cancer treatment groups (Table 3). Model effect estimates for the age 40+ Cox OC mortality model are given in Figure 3. Increased age, diabetes, hypertension, current smoking, previous stroke, and non-normal BMI all had harmful effects on OC mortality, while increased education and being married were protective. All model main effects were significant, and interactions were included between age and diabetes, age and education, age and hypertension, and age and previous stroke.

Calibration of the cause-specific model is given in Figure 5, while calibration of the Fine and Gray model refit in PLCO is given in Figure 4. Calibration of the Fine and Gray model refit in PLCO is excellent; the cause-specific model shows slight pessimism, particularly for patients at high risk of OC mortality.

*Other Cause Mortality Risk Drivers*

Predicted cumulative incidence curves from the refit Fine and Gray regression for 5 sample patients from PLCO are given in Figure 6. These examples demonstrate the sometimes nonintuitive ways that OC mortality operates. Age is clearly a major driver (the lowest risk patient is 58 while the highest risk patient is 80); however, comorbidities also clearly play a role. The three highest risk patients are all around 80; the highest risk patient’s OC mortality is driven by his underweight BMI, while the moderate and high risk patients’ risk appears driven by their marital status and lower educational attainment. OC mortality predictions can be further explored using the app given at https://elizabethchase.shinyapps.io/pctreat/.

These insights into the drivers of OC mortality are further explored in Table 4. Older patients are at higher risk of OC mortality, while high risk patients are more likely to be black and other race. Graduating from college is highly protective for OC mortality, as is being married. Current smokers and former smokers are overrepresented in the high and very high OC mortality risk groups, as are underweight and obese patients. Despite not being considered in modeling, prostate cancer characteristics were also linked with OC mortality predictions. Patients with more aggressive cancers were overrepresented in the high and very high OC mortality risk groups, while patients with less aggressive cancers generally had reduced OC mortality risk. Patients in the high and very high OC mortality risk groups were also more likely to die of their prostate cancer. Sample sizes for aggressive cancers are quite low, though, given the nature of the PLCO trial, so conclusions about these patients must be made with care.

*Treatment Concordance*

Figure 8 shows the relationship between prostate cancer mortality risk, OC mortality risk, and treatment assignment, with each dot representing a patient and the color giving his primary treatment. Clinicians seem to do well at identifying patients at very high and high risk of OC mortality, and most of these patients did not receive curative therapy. Similarly, clinicians seem to successfully identify patients at very low OC mortality risk and prescribe them more aggressive treatment—even when their prostate cancer characteristics might not warrant it. However, clinicians appear to struggle to accurately classify men in the low and moderate OC mortality risk categories. Many of the men at moderate OC mortality risk and unfavorable intermediate/high/very high PC mortality risk are being placed on active surveillance, even though most of these men should receive curative treatment; the same is true for the low OC mortality risk men with favorable intermediate PC risk who are placed on active surveillance.

Using our OC mortality prediction model and information on PLCO patients’ prostate cancers, we made treatment recommendations according to the 2020 NCCN guidelines [2]. We compared these recommendations with the treatments the men actually received, and classify them as concordant, too aggressive, or not aggressive enough in Table 4. We see that roughly a quarter of very low OC mortality risk patients are receiving treatment that is too aggressive, while roughly half of moderate and high risk OC mortality risk patients are receiving treatment that is nonconcordant with recommendations. For high risk patients, 30.6% of patients receive treatment that is not aggressive enough, suggesting excessive pessimism in estimating OC mortality risk; 22.2% receive treatment that is too aggressive.

**Discussion**

We present a parsimonious and accurate prediction model for OC mortality in US prostate cancer patients, built and tested in two diverse, prospective national cohorts. Our model shows excellent discrimination and calibration despite being validated in a population that is very different from its training population. It requires only eight easily obtained predictors and can be accessed at https://elizabethchase.shinyapps.io/pctreat/. We believe it will be of particular use for patients at low to moderate risk of OC mortality with favorable intermediate to high risk prostate cancer. It joins a growing list of OC mortality prediction models for men with prostate cancer; however, our model stands out for its simplicity, usability and impressive predictive performance in diverse patient populations.

When building the model, we considered a long list of potential predictors and several modeling strategies. The statistical methods used to produce the model are rigorous and do not rely on intermediate risk scores, and we present risk predictions over time, rather than point estimates. These features make it very attractive for future statistical work, and by providing all training data and code used to produce the model, our model can be combined with other prediction tools and treatment effect estimates to build new models. We have also made it easy for future researchers to reproduce our analysis and perform similar analyses for new disease sites and applications. This is particularly important because the approach we use here—building an OC mortality 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, and using cohorts likes NHANES vastly increases the number and quality of the OC mortality predictors that can be considered as candidates. Further research is needed, but this idea may be useful to researchers working on breast cancer, head and neck cancer, and many other disease sites where OC mortality is the primary driver of mortality.

Our study is subject to several limitations. Although it has its strengths, building a cancer OC mortality 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, so we believe the risk of bias is low. However, this approach limits our ability to obtain predictions of OC cumulative incidence without refitting the model within PLCO (what we do here) or combining our model with other prostate cancer specific mortality prediction tools. Our decision to not use the NHANES survey weights is another weakness of our analysis. If used properly, the survey weights would make our final model nationally representative of the US population and would also correct somewhat for missingness. These strengths are formidable; however, the methodological challenges of incorporating the weights within a complex model comparison framework were too great. Future research—perhaps aided by sophisticated survey statisticians—will hopefully correct this shortcoming. National representativeness aside, though, NHANES remains a remarkably diverse sample, even when unweighted.

*Conclusions*

We have built and validated an accurate and succinct model for OC mortality in US prostate cancer patients that is ready for clinical use. Our model will improve quality of care when patient life expectancy is a factor in treatment decision-making. The generalizability of our approach and transparency of our documentation will aid future research on cancer OC mortality across disease sites.

**References**

1. *American Cancer Society*. 2019. Key Statistics for Prostate Cancer. https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html.
2. National Comprehensive Cancer Network. 2020. NCCN Guidelines Version 2.2020: Prostate Cancer.
3. JD Sammon, F Abdollah, A D'Amico, M Gettman, A Haese, N Suardi, A Vickers, QD Trinh. 2015. Predicting life expectancy in men diagnosed with prostate cancer. *European Association of Urology* 68: 756-65.
4. C Jeldres. 2012. Life expectancy estimation in prostate cancer patients. *Canadian Urological Association Journal* 6.5: 374-5.
5. KMYB Leung, WM Hopman, J Kawakami. 2012. Challenging the 10-year rule: the accuracy of patient life expectancy predictions by physicians in relation to prostate cancer management. *Canadian Urological Association Journal* 6.5: 367-73.
6. J Walz, A Gallina, P Perrotte, C Jeldres, QD Trinh, GC Hutterer, M Traumann, A Ramirez, SF Shariat, M McCormack, JP Perreault, F Bénard, L Valiquette, F Saad, PI Karakiewicz. 2007. Clinicians are poor raters of life-expectancy before radical prostatectomy or definitive radiotherapy for localized prostate cancer. *BJU International* 100.6.
7. TJ Daskivich, K Chamie, L Kwan, J Labo, R Palvolgyi, A Dash, S Greenfield, MS Litwin. 2010. Overtreatment of men with low-risk prostate cancer and significant comorbidity. *Cancer* 117.10: 2058-66.
8. RM Hoffman, T Koyama, PC Albertsen, MJ Barry, TJ Daskivich, M Goodman, AS Hamilton, JL Stanford, AM Stroup, AL Potosky, DF Penson. 2015. Self-reported health status predicts other-cause mortality in men with localized prostate cancer: results from the prostate cancer outcomes study. *Journal of General Internal Medicine* 30.7: 924-34.
9. TJ Daskivich, L Kwan, A Dash, C Saigal, MS Litwin. 2015. An age adjusted comorbidity index to predict long-term, other cause mortality in men with prostate cancer. *Journal of Urology* 194: 73-78.
10. M Kent, DF Penson, PC Albertsen, M Goodman, AS Hamilton, JL Stanford, AM Stroup, B Ehdaie, PT Scardino, AJ Vickers. 2016. Successful external validation of a model to predict other cause mortality in localized prostate cancer. *BMC Medicine* 14.25.
11. P Riviere, C Tokeshi, J Hou, V Nalawande, R Sarkar, AJ Paravat, M Schiaffino, B Rose, R Xu, JD Murphy. 2019. Claims-Based Approach to Predict Cause-Specific Survival in Men With Prostate Cancer. *JCO Clinical Cancer Informatics* 3: 1-7.
12. A Tan, YF Kuo, JS Goodwin. 2013. Predicting life expectancy for community-dwelling older adults from Medicare claims data. *American Journal of Epidemiology* 178.6: 974-83.
13. National Center for Health Statistics, Centers for Disease Control and Prevention. 2019. National Health and Nutrition Examination Survey Data, 1999-2010. https://wwwn.cdc.gov/nchs/nhanes/ContinuousNhanes/Default.aspx
14. DR Cox. 1972. Regression models and life tables. *Journal of the Royal Statistical Society Series B (Methodological)* 34.2: 187-220.
15. H Ishwaran, UB Kogalur, EH Blackstone, MS Lauer. 2008. Random survival forests. *Annals of Applied Statistics* 2.3: 841-60.
16. P Royston, MKB Parmar. 2002. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine* 21: 2175-97.
17. R Core Team, R Foundation for Statistical Computing. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria. https://www.R-project.org/
18. Division of Cancer Prevention, National Cancer Institute. 2019. Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO). https://prevention.cancer.gov/ major-programs/prostate-lung-colorectal-and-ovarian-cancer-screening-trial

**Supplementary Materials**

*Variable Definitions*

We constructed some of our variables from other NHANES variables. In particular:

1. We defined a patient as having hypertension if they reported a previous diagnosis of hypertension or if their blood pressure reading was hypertensive (we followed standard NHANES guidelines on constructing the blood pressure reading).
2. We defined a patient as having diabetes if they reported a previous diagnosis of diabetes or if their blood glucose met clinical criteria for diabetes.
3. We defined a patient as having high cholesterol if they reported high cholesterol, or if their blood work met clinical criteria for high cholesterol.
4. We defined patients as current smokers if they had smoked 100 cigarettes in their lifetime and reported smoking cigarettes every day or some days at present; we defined patients as former smokers if they had smoked 100 cigarettes in their lifetime and reported smoking not at all at present; we defined patients as never smokers if they had not smoked 100 cigarettes in their lifetime.
5. We defined being overweight as having BMI greater than or equal to 25 and less than 40 and being obese as having BMI greater than or equal to 40.
6. We collapsed marital status into three categories: married if the patient was currently married or living with a partner, separated if the patient was widowed, divorced, or separated from their partner, and single if the patient had never married.
7. We collapsed race into non-Hispanic white, non-Hispanic black, and all other races.

For all other variables, we used unmodified results from the NHANES questionnaires about demographics, health insurance, health care access, mental health, and medical history. For more information on variable definition and construction, please see the repository posted at: https://github.com/elizabethchase/PCOtherCause.