

## ORIGINAL RESEARCH

# Predicting Life Expectancy to Target Cancer Screening Using Electronic Health Record Clinical Data

Alexandra K. Lee, PhD, MSPH<sup>1,2</sup> , Bocheng Jing, MS<sup>2,3</sup>, Sun Y. Jeon, PhD<sup>1,2</sup>,  
W. John Boscardin, PhD<sup>1,2,4</sup>, and Sei J. Lee, MD, MAS<sup>1,2</sup>

<sup>1</sup>Division of Geriatrics, University of California, 4150 Clement St, VA181G, San Francisco, CA, USA; <sup>2</sup>San Francisco Veterans Affairs Medical Center, San Francisco, USA; <sup>3</sup>Northern California Institute for Research and Education, San Francisco, USA; <sup>4</sup>Division of Biostatistics, University of California, San Francisco, San Francisco, USA.

**BACKGROUND:** Guidelines recommend breast and colorectal cancer screening for older adults with a life expectancy >10 years. Most mortality indexes require clinician data entry, presenting a barrier for routine use in care. Electronic health records (EHR) are a rich clinical data source that could be used to create individualized life expectancy predictions to identify patients for cancer screening without data entry.

**OBJECTIVE:** To develop and internally validate a life expectancy calculator from structured EHR data.

**DESIGN:** Retrospective cohort study using national Veteran's Affairs (VA) EHR databases.

**PATIENTS:** Veterans aged 50+ with a primary care visit during 2005.

**MAIN MEASURES:** We assessed demographics, diseases, medications, laboratory results, healthcare utilization, and vital signs 1 year prior to the index visit. Mortality follow-up was complete through 2017. Using the development cohort (80% sample), we used LASSO Cox regression to select ~100 predictors from 913 EHR data elements. In the validation cohort (remaining 20% sample), we calculated the integrated area under the curve (iAUC) and evaluated calibration.

**KEY RESULTS:** In 3,705,122 patients, the mean age was 68 years and the majority were male (97%) and white (85%); nearly half (49%) died. The life expectancy calculator included 93 predictors; age and gender most strongly contributed to discrimination; diseases also contributed significantly while vital signs were negligible. The iAUC was 0.816 (95% confidence interval, 0.815, 0.817) with good calibration.

**CONCLUSIONS:** We developed a life expectancy calculator using VA EHR data with excellent discrimination and calibration. Automated life expectancy prediction using EHR data may improve guideline-concordant breast and colorectal cancer screening by identifying patients with a life expectancy >10 years.

J Gen Intern Med 37(3):499–506

DOI: 10.1007/s11606-021-07018-7

© Society of General Internal Medicine 2021

Received November 9, 2020

Accepted June 30, 2021

Published online July 29, 2021



Life expectancy is an underused<sup>1,2</sup> but critically important piece of information for medical decision-making.<sup>3,4</sup> Numerous clinical guidelines recommend incorporating life expectancy into cancer screening decisions<sup>5,6</sup> since screening exposes patients to burdens and risk immediately while the benefits of decreased cancer mortality occur 10 years later.<sup>7–10</sup> Screening patients who have <10 years of life expectancy exposes them to the potential harms of screening with little chance of benefit.<sup>6,10</sup> To increase the rates of guideline-concordant cancer screening, there is a clear need for accurate, easy-to-use life expectancy calculators to appropriately target cancer screening.

Existing methods for life expectancy estimation have clear limitations that impair their utility to guide screening decisions. Clinicians often rely on their training and intuition to estimate life expectancy, but studies have shown that clinicians frequently overestimate life expectancy.<sup>11</sup> Objective, data-driven mortality prediction indexes have been developed and are available at websites such as [ePrognosis.com](https://www.ePrognosis.com).<sup>12–18</sup> However, these mortality indexes are underused,<sup>1–3</sup> likely due in part to the substantial time burden needed for collecting and inputting patient data into an online calculator.<sup>19</sup>

A life expectancy calculator embedded within the EHR would obviate the need for clinicians' data entry to determine prognosis. While current clinical alerts are triggered by age, this EHR-embedded life expectancy calculator could suppress cancer screening reminders for patients with a limited life expectancy regardless of age and trigger cancer screening reminders for older patients with life expectancy ≥10 years.<sup>5</sup> Existing mortality prediction indexes include self-reported physical functioning, which is not routinely captured in the EHR system. Thus, developing new EHR-based life expectancy calculators is a critical first step to make it easier for clinicians to provide guideline-concordant care.

As the largest vertically integrated healthcare system in the US, the Veterans Affairs (VA) is an ideal setting to develop EHR-based life expectancy calculators to better target CRC screening. First, the VA has had a longstanding focus on preventive care.<sup>20</sup> Second, studies have found high rates of colorectal cancer screening among Veterans with high comorbidity burden despite the potential for harm in this population,<sup>21</sup> suggesting that targeting colorectal cancer screening

through the use of life expectancy may be especially important within VA. Finally, VA's EHR provides an ideal setting for life expectancy calculator development, with a large population with >10 years of follow-up and a wide variety of clinical data elements, including pharmacy, laboratory, and vital sign data.

Our objective in this study was to use the Veterans Affairs (VA) EHR database to build a life expectancy calculator for adults aged  $\geq 50$  years. We sought to utilize EHR structured data elements to develop a model with approximately 100 predictors to accurately estimate life expectancy.

## METHODS

### Study Population

We used the VA EHR databases to identify all Veterans aged  $\geq 50$  using VA primary care in 2005, with mortality follow-up through 2017. The first primary care visit during 2005 was designated as the index visit for each patient; we then used a 1-year "look-back" period prior to the index visit to capture six domains of predictor variables: demographics, disease diagnoses, medication use, laboratory results, vital signs, and healthcare utilization. Of the 3,971,825 VA patients aged  $\geq 50$  with outpatient visits during 2005, 3,708,465 had at least one primary care visit, 3,343 (0.09%) had an invalid date of death prior to the index primary care visit, resulting in a final sample size of 3,705,122. We separated our data into an 80% random sample for model development ( $n=2,964,098$ ) and the remaining 20% for model validation ( $n=741,024$ ).

### Predictor Assessment

For disease diagnoses, we collected all ICD-9 diagnosis codes from the inpatient and outpatient medical SAS files and applied the Clinical Classifications Software (CCS)<sup>22</sup> to create 281 diagnosis groups. The inpatient and outpatient medical SAS files are constructed by compiling patient-level EHR data from 170 local VA Medical Center EHR systems to a centralized data warehouse.<sup>23,24</sup> For medication use, we used the Veteran Health Administration's National Drug File (NDF) to identify 401 unique medication classes from the Pharmacy Benefits Management file and classified use as any or none in the past year. For each of the 71 laboratory tests, we created a five-level variable: normal, abnormal low, abnormal high, nonsensical values, or not measured. Data on seven vital signs were extracted from the inpatient and outpatient medical SAS files: pulse, temperature, systolic blood pressure, respiration, pain level, BMI (weight and height), and pulse oximetry. In addition, because weight changes can be strongly indicative of poor health,<sup>25–28</sup> we a priori calculated two variables: weight range as the largest difference between weights recorded in the year prior to the index date, and recent weight change as the difference between the most recent weight and the weight closest to two weeks prior to the index date. For healthcare

utilization, we created 149 types of healthcare visits according to visit stop code and categorized visit frequency as zero, one, or at least two visits. Demographic variables (age, gender, and race/ethnicity) were extracted from the Clinical Data Warehouse. Race/ethnicity was included in descriptive characteristics but was not incorporated into model building. We did not want the observed shorter life expectancies of minority groups<sup>29,30</sup> to influence treatment decisions that could potentially exacerbate existing racial/ethnic inequalities.<sup>31</sup> In total, we obtained 913 predictors from the EHR to use for model-building.

For missing demographics (other than age) and vital signs (other than pulse oximetry and pain), we conducted a single stochastic conditional mean imputation, using a regression equation with all variables with any missingness included. Variables with missingness were gender ( $n=40,615$ ; 1.1%), race/ethnicity ( $n=233,221$ ; 6.3%), pulse ( $n=221,574$ ; 6.0%), temperature ( $n=518,032$ ; 14.0%), respiration ( $n=449,069$ ; 12.1%), weight range ( $n=967,463$ ; 26.1%), recent weight change (978,289; 26.4%), systolic blood pressure ( $n=215,821$ ; 5.8%), and BMI (626,569; 16.9%).

Prior to modeling, we considered the functional form of predictor variables. We used a restricted cubic spline with 4 knots for age. We categorized BMI as <18.5, 18.5–25, 25–30, 30–35, >35, systolic blood pressure as <90, 90–104, 105–119, 120–139, 140–159, 160–179,  $\geq 180$ , and pulse oximetry as <93, 93–95,  $\geq 96$ , or no measurement, following commonly used clinical thresholds.

### Outcome Assessment

We used the Veterans Health Administration's Vital Status File<sup>32</sup> to identify all deaths through December 31, 2017, an approach with high accuracy compared to the National Death Index (98% sensitivity and 98% specificity).<sup>33</sup>

### Statistical Analysis

We first examined baseline characteristics of the development and validation cohorts. Using the development cohort ( $n=2,964,098$ ), we built our life expectancy prediction model in two phases: variable selection and model coefficient estimation. Two phases were needed because current statistical software does not allow least absolute shrinkage and selection operator (LASSO) regression in parametric survival models.

In the first variable selection phase of model development, we applied a LASSO Cox proportional hazards regression with a BIC-optimized lambda. To determine this lambda, we evaluated a sequence of 100 lambda values and chose the lambda that gave us the lowest BIC with <100 predictors. We sought to build a model with approximately 100 predictors after considering several factors. First, while there is little penalty for including additional predictors in EHR models (since additional data collection or data entry is not required), models with greater number of predictors are more likely to be overfit,<sup>34</sup> leading to reduced generalizability to different time

periods or slightly different patient populations. In addition, we found little difference in discrimination between models with all 913 predictors and 80 predictors. Thus, a ~100 predictor model appears to best balance optimizing model discrimination with maximizing model generalizability.

In the second coefficient estimation phase of model development, we used the variables selected from LASSO Cox regression and fit a Gompertz survival model in the development cohort ( $n=2,964,098$ ). We chose a parametric survival model to allow for life expectancy predictions beyond the observed follow-up time. Of the numerous parametric survival distributions, Gompertz is most appropriate for human mortality curves because it fits the observed doubling in mortality rate with each decade of life.<sup>35</sup>

To internally validate our model in the validation cohort ( $n=741,024$ ), we calculated the integrated area under the curve (iAUC),<sup>36,37</sup> which is a weighted average of the areas under the time-varying receiver operator characteristic curves at all possible cutpoints. We graphed the iAUC to display the AUC over time and at three time points of interest: 2 years, 5 years, and 10 years. We evaluated model calibration by creating ventiles (20 bins) of predicted life expectancy and comparing the mean of the predicted 10th percentile of life expectancy within each group to the observed 10th percentile of life expectancy within each group (time to 10% mortality).

To identify which domains of predictors (demographics, diseases, medications, labs, vital signs, and healthcare utilization) most strongly contributed to discrimination in the final model, we evaluated a series of models. First, we fit a Gompertz model with each domain of predictors alone, without any other predictors. Since demographics are known to have strong predictive power, we next fit models with demographics plus each domain of predictors separately. Finally, we evaluated the effect of removing one domain of predictors from the full model, fitting models with 5 of 6 domains. We fit all models in the development cohort and calculated the iAUC in the validation cohort.

To provide an example of the output a clinician would see from this life expectancy calculator, we generated two test cases, “Mr. Unhealthy” and “Mr. Healthy,” taken from individuals at the 10th and 80th percentiles of predicted median life expectancy, respectively. For each test case, we presented the 25th, 50th, and 75th percentile of predicted life expectancy, as well as the predicted probability of mortality within 2 years, 5 years, and 10 years. To provide an interpretation, the 25th percentile of life expectancy means that the individual has a 25% chance of dying before that time, while the median life expectancy is the time at which an individual has a 50/50 chance of being alive. All 95% confidence intervals for life expectancy estimates were  $\leq 0.1$  year and thus were not presented.

We conducted sensitivity analyses to ensure results were similar across clinically important subgroups. We examined iAUC and calibration plots across subgroups defined by age, gender, race/ethnicity, and presence of cancer.

## RESULTS

There were 2,964,098 patients in the development cohort and 741,024 patients in the validation cohort. Baseline characteristics were essentially the same in the two cohorts (Table 1). The mean age was 68.0 years, 97% were male, 85% were white, and 11% were black. The mean BMI was 28.9. Over half of patients had  $\geq 8$  diseases and 42% had  $\geq 6$  medications. During a median of 12.1 years of follow-up, 49% of patients died (1,453,767 in the development cohort and 362,934 in the validation cohort).

The final life expectancy calculator had 93 predictors from six domains: 2 demographic predictors (age and sex), 39 disease predictors, 28 medication predictors, 14 laboratory predictors, 4 vital sign predictors, and 6 healthcare utilization predictors (Supplemental Table S1). The iAUC was 0.816 in the development cohort and 0.816 in the validation cohort (95% confidence interval in the validation cohort: 0.815, 0.817). The AUC was consistently around 0.8 throughout follow-up time in the validation cohort (Figure 1), with the lowest value of 0.803 at 2 years and increasing to 0.809 at 5 years and 0.830 at 10 years.

Several predictors had a counterintuitive association with mortality. For example, osteoarthritis, benign neoplasms, and upper respiratory infections were all associated with decreased mortality risk and longer life expectancy. To determine the relative importance of these counterintuitive factors, we conducted sensitivity analyses examining the discrimination of models with these 3 factors removed and found that removing these factors led to trivial declines in iAUC (0.8162 full model, 0.8156 model without 3 counterintuitive factors) (Supplemental Table S2). Since our primary goal was to maximize discrimination and previous prediction studies have also shown counterintuitive factors,<sup>38</sup> we elected to retain our counterintuitive factors in our prediction model.

The model was well calibrated in the validation cohort. Across 20 bins of predicted life expectancy, the observed time to 10% mortality aligned with the predicted 10th percentile of life expectancy (Figure 2). For example, in the 15th ventile, the predicted 10th percentile of life expectancy was 5.2 years and the observed time to 10% mortality was 5.5 years. For individuals with predicted median life expectancy  $< 12$  years, predicted and observed time to 50% mortality was well aligned. In the 6th ventile, the predicted median life expectancy was 9.2 years and the observed time to 50% mortality was 8.6 years.

In subgroups defined by age, gender, race/ethnicity, and presence of cancer, the life expectancy calculator showed good discrimination and calibration (Supplemental Table S3). Across age groups, discrimination was good but notably lower than the overall sample, likely due to strong predictive power of age (range of iAUC stratified on age: 0.743–0.776). Discrimination was very similar by race and presence of cancer, while discrimination was better in women (iAUC: 0.860) than men (iAUC: 0.814). Calibration was

Table 1 Baseline Characteristics of Veteran's Affairs Primary Care Patients Aged 50+ in 2005, n=3,705,122

Characteristic	Development cohort N=2,964,098	Validation cohort N=741,024
Demographics		
Age, mean $\pm$ SD	68.0 $\pm$ 10.4	68.0 $\pm$ 10.4
Male	2,822,115 (96.5)	714,705 (96.5)
Race/Ethnicity		
Non-Hispanic White	2,537,088 (85.3)	631,848 (85.3)
Non-Hispanic Black	331,548 (11.2)	83,058 (11.2)
Other	95,462 (3.2)	26,118 (3.5)
Vital signs		
BMI, mean $\pm$ SD	28.9 $\pm$ 5.5	28.9 $\pm$ 5.5
Systolic blood pressure, mean $\pm$ SD	134.7 $\pm$ 18.6	134.6 $\pm$ 18.6
Laboratory values*		
Albumin	4.1 [3.8, 4.3]	4.1 [3.8, 4.3]
Serum creatinine	1.1 [0.9, 1.3]	1.1 [0.9, 1.3]
LDL, mean $\pm$ SD	104.5 $\pm$ 35.1	104.5 $\pm$ 35.1
HbA1c, mean $\pm$ SD	6.8 $\pm$ 1.6	6.8 $\pm$ 1.6
Diseases (per person)		
0-1	131,696 (4.4)	32,830 (4.4)
2-4	528,927 (17.8)	132,306 (17.9)
5-7	808,827 (27.3)	202,364 (27.3)
8+	1,494,648 (50.4)	373,524 (50.4)
Medications (per person)		
0	481,438 (16.2)	120,313 (16.2)
1-2	434,173 (14.7)	108,295 (14.6)
3-5	812,765 (27.4)	203,496 (27.5)
6+	1,235,722 (41.7)	308,920 (41.7)
Healthcare utilization		
Number of hospitalizations, median [25th, 75th percentiles]	0 [0, 0]	0 [0, 0]
Number of emergency department, visits, median [25th, 75th percentiles]	0 [0, 0]	0 [0, 0]
Number of outpatient visits, median [25th, 75th percentiles]	2 [1, 4]	2 [1, 4]
Died within 2 years	233,626 (7.9)	57,840 (7.8)
Died within 5 years	600,316 (20.3)	150,061 (20.3)
Died within 10 years	1,200,848 (40.5)	300,048 (40.5)

Values are n (%) unless noted as mean  $\pm$  SD, or median [25th, 75th percentiles]

\*Laboratory values were categorized as normal, abnormal low, abnormal high, nonsensical values, or not measured. Albumin was not measured in 44% of both the development and validation cohorts. Serum creatinine was not measured in 30% of the development and validation cohorts. LDL was not measured in 44% of the development and validation cohorts. HbA1c was not measured in 72% of the development and validation cohorts

excellent by gender, race/ethnicity, and presence of cancer (Supplemental Figure S1); within age groups, calibration was excellent for ages  $\geq 65$  and good for ages 50–64.

As expected, diseases and demographics contributed highly to discrimination (Figure 3). Including only age and gender produced a model with an iAUC of 0.710, including only diseases produced an iAUC of 0.725, while including only healthcare utilization produced an iAUC of 0.516. When adding to a demographics-only model, the iAUC was highest when including diseases (0.790). Removing demographics

from the full model had the largest impact (iAUC drop of 0.0403 to 0.7760), while removing healthcare utilization had a negligible impact (iAUC drop of 0.0007 to 0.8156).

In the example output for “Mr. Healthy” (Table 2), predicted 25th percentile life expectancy was 13.7 years and predicted median life expectancy was 20.7 years, indicating that cancer screening would be recommended. In contrast, Mr. Unhealthy’s predicted median life expectancy is only 5.4 years, and his 75th percentile life expectancy is 8.8 years, indicating that cancer screening would be more likely to harm than help.

## DISCUSSION

We developed and internally validated a life expectancy prediction calculator using structured data from the VA EHR; our final model included 93 predictors and had an iAUC of 0.816. As expected, diseases and demographics were the strongest contributors to high discrimination. Our model demonstrated good calibration overall and across demographic subgroups.

Our life expectancy calculator performs comparably to other long-term mortality risk tools that had C-statistics around 0.8,<sup>12–14,16,39</sup> but has two notable differences. First, our life

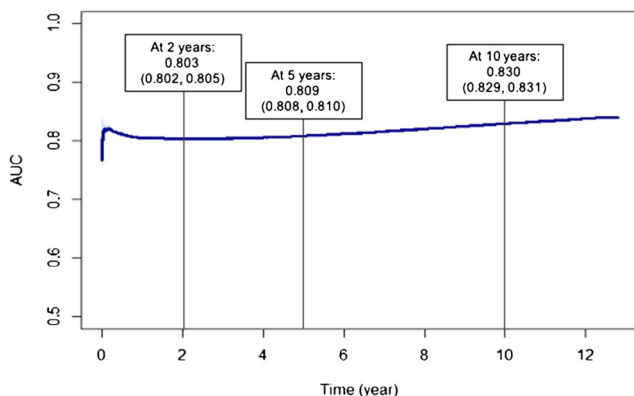
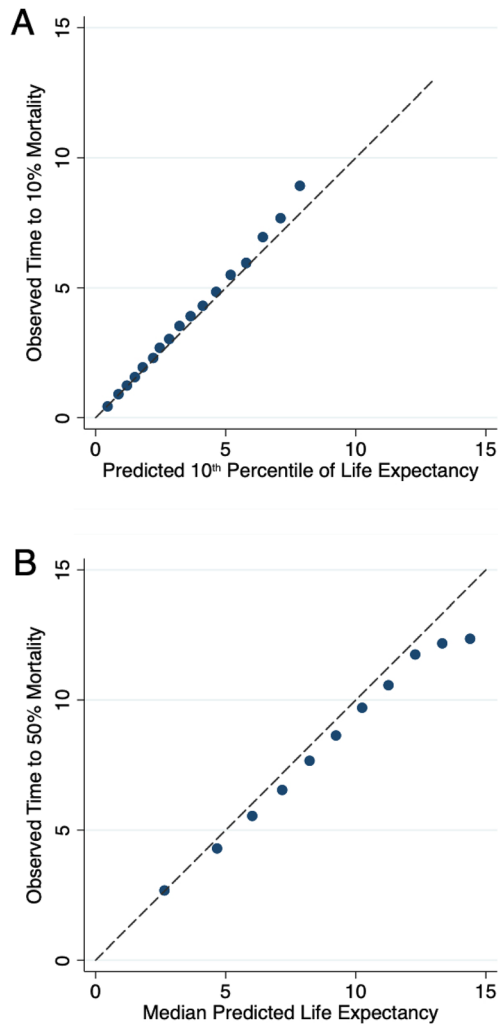


Figure 1. AUC over time from life expectancy calculator. overall, the iAUC was 0.8162 (0.8153, 0.8171)



**Figure 2.** Calibration plot of predicted vs. observed remaining life expectancy. Each point represents a ventile (1/20) of the validation cohort, representing 37,051 patients. In panel A, 3 lowest-risk ventiles could not be shown since <10% of patients in those ventiles had died by the end of follow-up. In panel B, 9 lowest-risk ventiles could not be shown since <50% of patients in those ventiles had died by the end of follow-up

expectancy calculator used EHR data while other mortality risk tools relied on self-reported data from questionnaires.<sup>12–16,40</sup> Using structured data elements from the EHR, we were able to consider nearly one thousand predictors rather than several dozen predictors considered in other models. Our life expectancy calculator included many more predictors than other indexes because it was designed to be embedded within the EHR, unlike previous mortality risk indexes that rely on self-reported data and clinician data entry. To our knowledge, only the Mathias mortality risk prediction equation used EHR data from a single multi-specialty group practice; they applied predictive mining to select 24 predictors from 980 possible predictors in approximately 8,000 patients.<sup>39</sup>

The second major difference is that our calculator predicts remaining life expectancy, while other indexes generally only predict mortality risk at one specific time point. Most mortality risk prediction tools are for short-term mortality ( $\leq 5$

years),<sup>39,40</sup> only a few predict  $\geq 10$ -year mortality risk, which is needed for cancer screening.<sup>13,16,18</sup> Of other indexes, only the Lee index predicts life expectancy.<sup>15</sup> Remaining life expectancy is difficult to predict because it requires extended follow-up time and relatively high cumulative mortality to generate accurate predictions. The benefit is that life expectancy is more intuitive for patients and provides a more complete picture than mortality risk at a single time.

Our model had better discrimination among females, who were approximately 3% of the study sample, compared to males. We believe this may be due to the fact that women in our VA population were substantially younger than men (mean age 63 vs. 68 for men). Previous studies have shown that measures of discrimination are higher for mortality indexes in younger populations compared to older populations.<sup>41</sup> Ultimately, it is reassuring that the model performs well in women, even though they were only 3% of the data, suggesting that our life expectancy predictions can be used to target breast cancer screening.

Consistent with other mortality prediction tools, we did not incorporate race/ethnicity into our life expectancy calculator.<sup>12,14,16</sup> While non-white individuals tend to have shorter life expectancies,<sup>29,30</sup> prediction algorithms incorporating race/ethnicity may perpetuate and exacerbate existing disparities and racism in the medical system.<sup>31</sup> However, excluding race/ethnicity from algorithms can still lead to racially biased models,<sup>42</sup> and machine learning approaches may be particularly prone to encode racial/ethnic bias in the healthcare system.<sup>43–45</sup> We took numerous steps to minimize potential racial bias, including manual variable cleaning and specification (e.g., assigning all laboratory values to one of five categories). We welcome the development of “algorithmic stewardship” that will advance the equitable development and use of prediction tools in clinical medicine.<sup>46,47</sup>

Uncertainty is inherent in the prediction of all future events, and we have quantified this uncertainty in the interquartile range for an individual’s predicted life expectancy. While predictions of “very high” risk or “almost no” risk are easier to understand, communicate, and act upon, many modern clinical decisions rely on modest risk of future events, including predicted cardiovascular risk for statin medications<sup>48</sup> and predicted fracture risk for osteoporosis medications.<sup>49</sup> We believe it is critical to acknowledge the uncertainty and to have this uncertainty inform the acceptable level of risk and possibility of benefit.<sup>50</sup> While some patients may want to make decisions based on the most likely scenario, others may want to continue with treatments even if they have a small chance of benefitting. Eliciting and clarifying these preferences is a core component of shared decision-making.<sup>51</sup> Explicitly noting the prediction uncertainty will enable nuanced discussions between clinicians and patients.

Our work has clear implications. Programming our life expectancy calculator into the VA EHR can improve guideline-concordant care for numerous preventive therapies with different times to benefit, acting at both the patient level

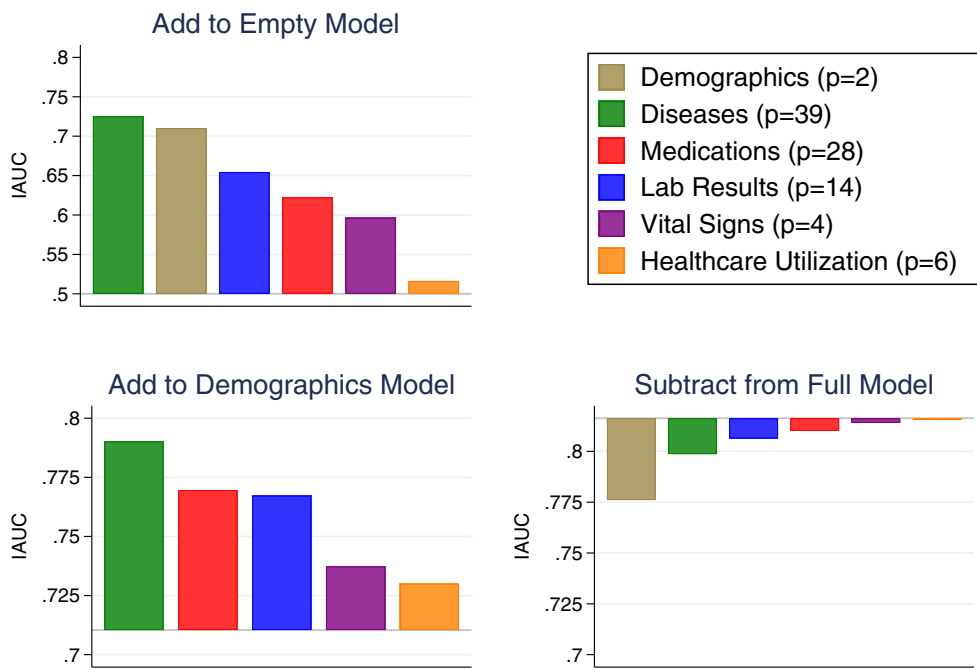


Figure 3 The contributions of different domains of variables to iAUC in the validation cohort

and at the system level. For an individual patient, the median life expectancy provides the best estimate of the remaining lifetime. From a systems perspective, integration into the EHR will enable widespread uptake as well as tailoring of clinical decision support systems. For example, clinical reminders for colorectal cancer screening can be suppressed for patients with a life expectancy <10 years instead of using age-based cutoffs, thereby reducing inappropriate screening among younger sicker patients. Similarly, clinical reminders could be triggered for adults who are outside of the usual age range for colorectal cancer screening but have >10 year life expectancy and thus may benefit from screening.<sup>52,53</sup> Using life expectancy rather than age will also help minimize unintended harms of quality measures.<sup>21</sup> This life expectancy calculator can be replicated in other EHR settings to improve guideline-concordant care beyond the VA.

Our study had noteworthy strengths. First, we were able to evaluate nearly 1000 predictor variables in 3.7 million Veterans. Second, we used state-of-the-art methods (LASSO Cox regression and Gompertz regression) to build our final model, after determining that machine learning methods would

provide not clinically significant improvements in model performance.<sup>54,55</sup> Finally, we demonstrated that our model was well-calibrated and had high discrimination in numerous subgroups.

Our study had limitations. First, the performance of our life expectancy calculator outside of the VA population is unclear; future studies should validate and/or develop new EHR-based prediction indexes in non-VA populations, particularly given that women made up only 3% of our sample. Our calculator can be translated to other EHR systems using the OMOP Common Data Model.<sup>56</sup> Second, we were not able to incorporate other risk factors that influence mortality, including social risk factors and physical function. Third, we were only able to incorporate healthcare utilization within the VA system, while some Veterans receive care at outside medical institutions. Fourth, life expectancies >12 years are model extrapolations to be interpreted with caution. However, these extrapolations will have little impact on clinical utility, since any life expectancy >10 years indicates likely to benefit from cancer screenings. Fifth, the presence of counterintuitive factors may concern some users and slow widespread adoption.

Table 2 Sample Output from Life Expectancy Prediction Calculator: Life Expectancy Prediction and Predicted Probability of Mortality

	Life expectancy prediction (years)			Predicted probability (%) of mortality		
	25th percentile life expectancy (time to 25% mortality risk)	Median life expectancy (time to 50% mortality risk)	75th percentile life expectancy (time to 75% mortality risk)	Within 2 years	Within 5 years	Within 10 years
Example 1: “Mr. Healthy”	13.7 years**	20.7 years**	26.6 years**	2.0%	5.9%	15.2%
Example 2: “Mr. Unhealthy”	2.6 years	5.4 years	8.8 years	19.1%	46.7%	81.7%

\*\*These data are beyond the observed follow-up time and are model extrapolations; use caution when interpreting

However, as “black box” prediction algorithms (i.e., neural networks for breast cancer histology)<sup>57</sup> become more common in clinical medicine, we believe that users will increasingly focus more on predictive accuracy and less on model transparency.<sup>58</sup> Finally, predictions may not be accurate for patients with rare diseases.

In conclusion, we developed and internally validated a life expectancy calculator with excellent discrimination and calibration using EHR data for Veterans aged  $\geq 50$  years. With excellent calibration across 1 to 10 years of mortality prediction, our life expectancy calculator could be useful for a wide variety of clinical decisions, such as statin treatment and glycemic control. Accurate estimates of life expectancy are critical for breast and colorectal cancer screening and building in such estimates into the EHR system will be a critical foundational step in improving the rates of individualized, guideline-concordant care.

**Corresponding Author:** Alexandra K. Lee, PhD, MSPH; Division of Geriatrics, University of California, 4150 Clement St, VA181G, San Francisco, CA 94121, USA (e-mail: alexandra.lee@ucsf.edu).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11606-021-07018-7>.

**Funding** This study was supported by the Department of Veteran's Affairs Office of Research and Development, Health Services Research and Development Investigator Initiated Research grant 5101HX002135. A.K.L. was supported by NIH/NIA grant T32AG000212.

#### Declarations:

**Disclaimer:** The content of this article does not represent the view of the VA or the U.S. government. This article represents the views of the authors and not those of the VA or the Health Services Research and Development.

**Conflict of Interest:** All authors declare no conflicts of interest.

## REFERENCES

- Schoenborn NL, Bowman TL, Cayea D, Pollack CE, Feeser S, Boyd C. Primary Care Practitioners' Views on Incorporating Long-term Prognosis in the Care of Older Adults. *JAMA Intern Med.* 2016;212(4):671-678. <https://doi.org/10.1001/jamainternmed.2016.0670>
- Thomas JM, Fried TR. Defining the Scope of Prognosis: Primary Care Clinicians' Perspectives on Predicting the Future Health of Older Adults. *J Pain Symptom Manage.* 2018;55(5):1269-1275.e1. <https://doi.org/10.1016/j.jpainsymman.2018.01.017>
- Gill TM. The central role of prognosis in clinical decision making. *JAMA.* 2012;307(2):199-200. <https://doi.org/10.1001/jama.2011.1992>
- Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: What, why, and how? *BMJ.* 2009;338:b375. <https://doi.org/10.1136/bmj.b375>
- Smith RA, Andrews KS, Brooks D, et al. Cancer screening in the United States, 2018: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J Clin.* 2018;68(4):297-316. <https://doi.org/10.3322/caac.21446>
- Wilt TJ, Harris RP, Qaseem A, et al. Screening for cancer: Advice for high-value care from the American college of physicians. *Ann Intern Med.* 2015;162(10):718-725. <https://doi.org/10.7326/M14-2326>
- Lee SJ, Leipzig RM, Walter LC. Incorporating lag time to benefit into prevention decisions for older adults. *JAMA.* 2013;310(24):2609-2610. <https://doi.org/10.1001/jama.2013.282612>
- Walter LC, Covinsky KE. Cancer Screening in Elderly Patients. *JAMA.* 2001;285(21):2750-2756. <https://doi.org/10.1001/jama.285.21.2750>
- Tang V, Boscardin WJ, Stijacic-Cenzer I, Lee SJ. Time to benefit for colorectal cancer screening: Survival meta-analysis of flexible sigmoidoscopy trials. *BMJ.* 2015;350:h1662. <https://doi.org/10.1136/bmj.h1662>
- Lee SJ, Boscardin WJ, Stijacic-Cenzer I, Conell-Price J, O'Brien S, Walter LC. Time lag to benefit after screening for breast and colorectal cancer: meta-analysis of survival data from the United States, Sweden, United Kingdom, and Denmark. *Br Med J.* 2013;346(3):e8441. <https://doi.org/10.1136/bmj.e8441>
- Christakis NA, Lamont EB. Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ.* 2000;320:469-473. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=27288&tool=pmcentrez&rendertype=abstract>
- Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. *Jama.* 2006;295(7):801-808. <https://doi.org/10.1001/jama.295.7.801>
- Cruz M, Covinsky K, Widera E, Stijacic-Cenzer I, Lee S. Predicting 10 year mortality for older adults. *J Am Med Assoc.* 2013;309(9):875-876. <https://doi.org/10.1136/bmj.f174.2>
- Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. *J Gen Intern Med.* 2009;24(10):1115-1122. <https://doi.org/10.1007/s11606-009-1073-y>
- Lee SJ, Boscardin WJ, Kirby KA, Covinsky KE. Individualizing life expectancy estimates for older adults using the gompertz law of human mortality. *PLoS One.* 2014;9(9):e108540. <https://doi.org/10.1371/journal.pone.0108540>
- Schonberg MA, Li V, Marcantonio ER, Davis RB, McCarthy EP. Predicting Mortality up to 14 Years Among Community-Dwelling Adults Aged 65 and Older. *J Am Geriatr Soc.* 2017;65(6):1310-1315. <https://doi.org/10.1111/jgs.14805>
- Hayes AJ, Leal J, Kelman CW, Clarke PM. Risk equations to predict life expectancy of people with Type2 diabetes mellitus following major complications: A study from Western Australia. *Diabet Med.* 2011;28(4):428-435. <https://doi.org/10.1111/j.1464-5491.2010.03189.x>
- Tan A, Kuo YF, Goodwin JS. Predicting life expectancy for community-dwelling older adults from Medicare claims data. *Am J Epidemiol.* 2013;178(6):974-983. <https://doi.org/10.1093/aje/kwt054>
- Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. *Int J Med Inform.* 2017;98:56-64. <https://doi.org/10.1016/j.ijmedinf.2016.12.001>
- Jha AK, Wright SM, Perlin JB. Performance measures, vaccinations, and pneumonia rates among high-risk patients in veterans administration health care. *Am J Public Health.* 2007;97(12):2167-2172. <https://doi.org/10.2105/AJPH.2006.099440>
- Saini SD, Vijan S, Schoenfeld P, Powell AA, Moser S, Kerr EA. Role of quality measurement in inappropriate use of screening for colorectal cancer: Retrospective cohort study. *BMJ.* 2014;348:g1247. <https://doi.org/10.1136/bmj.g1247>
- Healthcare Cost and Utilization Project. Clinical Classifications Software (CCS) for ICD-9-CM Fact Sheet. Agency for Healthcare Research and Quality. <https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp>. Published 2012.
- (VIRc) VIRc. *VIRc Research User Guide: Fiscal Year 2014 VHA Medical SAS Outpatient Datasets and Inpatient Encounters Dataset*. Hines, IL; 2015.
- (VIRc) VIRc. *VIRc Research User Guide: Fiscal Year 2014 VHA Medical SAS Inpatient Datasets*. Hines, IL; 2015.
- Wannamethee SG, Shaper AG, Walker M. Weight Change, Weight Fluctuation, and Mortality. *Arch Intern Med.* 2002;162:2575-2580.
- Wannamethee SG, Shaper AG, Lennon L. Reasons for intentional weight loss, unintentional weight loss, and mortality in older men. *Arch Intern Med.* 2005;165(9):1035-1040. <https://doi.org/10.1001/archinte.165.9.1035>
- Wedick NM, Barrett-connor E, Knoke JD, Wingard DL. The relationship between weight loss and all-cause mortality in older men and women

- with and without diabetes mellitus: the Rancho Bernardo Study. *J Am Geriatr Soc.* 2002;50:1810-1815.
28. Fried LP, Tangen CM, Walston J, et al. Frailty in Older Adults: Evidence for a Phenotype. *J Gerontol.* 2001;56(3):146-156. <https://doi.org/10.1093/gerona/56.3.M146>
  29. Olshansky SJ, Antonucci T, Berkman L, et al. Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Aff.* 2012;31(8):1803-1813. <https://doi.org/10.1377/hlthaff.2011.0746>
  30. Harper S, Rushani D, Kaufman JS. Trends in the black-white life expectancy gap, 2003-2008. *JAMA.* 2012;307(21):2257-2259. <https://doi.org/10.1001/jama.2012.5059>
  31. Vyas DA, Eisenstein LG, Jones DS. Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms. Malina D, ed. *N Engl J Med.* 2020;383(9):874-882. <https://doi.org/10.1056/NEJMms2004740>
  32. Maynard C. Ascertaining Veterans' Vital Status: VA Data Sources for Mortality Ascertainment and Cause of Death Database & Methods Cyberseminar Series. VA Information Resource Center. [https://www.hsrd.research.va.gov/for\\_researchers/cyber\\_seminars/archives/1242-notes.pdf](https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/1242-notes.pdf). Published 2017. Accessed July 17, 2020.
  33. Sohn MW, Arnold N, Maynard C, Hynes DM. Accuracy and completeness of mortality data in the Department of Veterans Affairs. *Popul Health Metr.* 2006;4:1-8. <https://doi.org/10.1186/1478-7954-4-2>
  34. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.* Springer; 2009.
  35. Olshansky SJ, Carnes BA. Ever since Gompertz. *Demography.* 1997;34(1):1-15. <http://www.ncbi.nlm.nih.gov/pubmed/9074828>
  36. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics.* 2000;56(2):337-344. <https://doi.org/10.1111/j.0006-341X.2000.00337.x>
  37. Guo C, So Y, Jang W. Evaluating Predictive Accuracy of Survival Models with PROC PHREG. *Proc SAS Glob Forum 2017 Conf.* 2017:1-16. <https://pdfs.semanticscholar.org/0f63/7c13f7eac0dbb1a691da46197593fa131b.pdf>
  38. Lezzoni LI, Foley SM, Daley J, Hughes J, Fisher ES, Heeren T. Comorbidities, Complications, and Coding Bias: Does the Number of Diagnosis Codes Matter in Predicting In-Hospital Mortality? *JAMA.* 1992;267(16):2197-2203. <https://doi.org/10.1001/jama.1992.03480160055034>
  39. Mathias JS, Agrawal A, Feinglass J, Cooper AJ, Baker DW, Choudhary A. Development of a 5 year life expectancy index in older adults using predictive mining of electronic health record data. *J Am Med Informatics Assoc.* 2013;20(E1):118-124. <https://doi.org/10.1136/amiajnl-2012-001360>
  40. Yourman LC, Lee SJ, Schonberg M a, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. *JAMA.* 2012;307(2):182-192. <https://doi.org/10.1001/jama.2011.1966>
  41. Lee SJ, Go AS, Lindquist K, Bertenthal D, Covinsky KE. Chronic conditions and mortality among the oldest old. *Am J Public Health.* 2008;98(7):1209-1214. <https://doi.org/10.2105/AJPH.2007.130955>
  42. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science (80-).* 2019;366(6464):447-453. <https://doi.org/10.1126/science.aax2342>
  43. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care - Addressing Ethical Challenges. *N Engl J Med.* 2018;378(11):981-983. <https://doi.org/10.1056/NEJMp1714229>
  44. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Intern Med.* 2018;178(11):1544-1547. <https://doi.org/10.1001/jamainternmed.2018.3763>
  45. Goodman SN, Goel S, Cullen MR. Machine learning, health disparities, and causal reasoning. *Ann Intern Med.* 2018;169(12):883-885. <https://doi.org/10.7326/M18-3297>
  46. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Ann Intern Med.* 2018;169(12):866-872. <https://doi.org/10.7326/M18-1990>
  47. Eaneff S, Obermeyer Z, Butte AJ. The Case for Algorithmic Stewardship for Artificial Intelligence and Machine Learning Technologies. *JAMA - J Am Med Assoc.* 2020;324(14):1397-1398. <https://doi.org/10.1001/jama.2020.9371>
  48. Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol.* 2019;73(24):e285-e350. <https://doi.org/10.1016/j.jacc.2018.11.003>
  49. Camacho PM, Petak SM, Binkley N, et al. American Association of Clinical Endocrinologists and American College of Endocrinology Clinical Practice Guidelines for the Diagnosis and Treatment of Postmenopausal Osteoporosis - 2016. *Endocr Pract.* 2016;22(Suppl 4):1-42. <https://doi.org/10.4158/EP161435.GL>
  50. Smith AK, White DB, Arnold RM. Uncertainty—the other side of prognosis. *N Engl J Med.* 2013;368(26):2448-2450. <https://doi.org/10.1056/NEJMp1303295>
  51. Elwyn G, Cochran N, Pignone M. Shared Decision Making—The Importance of Diagnosing Preferences. *JAMA Intern Med.* 2017;177(9):1239-1240. <https://doi.org/10.1001/jamainternmed.2017.1923>
  52. Walter LC, Lindquist K, Nugent S, et al. Impact of age and comorbidity on colorectal cancer screening among older veterans. *Ann Intern Med.* 2009;150(7):465-473. <https://doi.org/10.7326/0003-4819-150-7-200904070-00006>
  53. Walter LC, Bertenthal D, Lindquist K, Konety BR. PSA screening among elderly men with limited life expectancies. *J Am Med Assoc.* 2006;296(19):2336-2342. <https://doi.org/10.1001/jama.296.19.2336>
  54. Jeon SY, Jing B, Lee AK, Boscardin WJ, Lee SJ. An Empirical Comparison of LASSO and Stepwise Selection Prediction Model Development Strategies with Various Sample Sizes. *Poster presented at the Society for Epidemiologic Research Conference,* Dec 2020.
  55. Jing B, Jeon SY, Lee AK, Boscardin WJ, Lee SJ. Comparing Traditional Regression Approaches with Machine Learning Approaches for 10-year Mortality Prediction. In: *HSR&D/QUERI National Conference.* Washington, DC; 2019.
  56. (OHDSI) OHDS and I. Data Standardization: OMOP Common Data Model. <https://www.ohdsi.org/data-standardization/>. Accessed February 16, 2021.
  57. Rakhlin A, Shvets A, Iglavikov V, Kalinin AA. Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. *arXiv.* 2018;1802.00752. [https://doi.org/10.1007/978-3-319-93000-8\\_83](https://doi.org/10.1007/978-3-319-93000-8_83)
  58. London AJ. Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability. *Hastings Cent Rep.* 2019;49(1):15-21. <https://doi.org/10.1002/hast.973>

**Publisher's Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.