QUERI

Life Expectancy Calculators

June 2016

Prepared for:
Department of Veterans Affairs
Veterans Health Administration
Quality Enhancement Research Initiative
Health Services Research & Development Service
Washington, DC 20420

Prepared by:
Evidence-based Synthesis Program (ESP) Center
Minneapolis VA Medical Center
Minneapolis, MN
Tim Wilt, MD, MPH, Director
Nancy Greer, PhD, Program Manager

Investigators:
Principal Investigator:
Thomas Rector, PhD, PharmD

Co-investigators:
Brent Taylor, PhD
Shahnaz Sultan, MD, MHSc
Aasma Shaukat, MD, MPH
Selcuk Adabag, MD, MS
David Nelson, PhD
Timothy Capecchi, MD

Research Associate:
Roderick MacDonald, MS

VA HEALTH CARE Defining EXCELLENCE in the 21st Century

PREFACE

The VA Evidence-based Synthesis Program (ESP) was established in 2007 to provide timely and accurate syntheses of targeted healthcare topics of particular importance to clinicians, managers, and policymakers as they work to improve the health and healthcare of Veterans. QUERI provides funding for four ESP Centers, and each Center has an active University affiliation. Center Directors are recognized leaders in the field of evidence synthesis with close ties to the AHRQ Evidence-based Practice Centers. The ESP is governed by a Steering Committee comprised of participants from VHA Policy, Program, and Operations Offices, VISN leadership, field-based investigators, and others as designated appropriate by QUERI/HSR&D.

The ESP Centers generate evidence syntheses on important clinical practice topics. These reports help:

- Develop clinical policies informed by evidence;
- Implement effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures; and
- Set the direction for future research to address gaps in clinical knowledge.

The ESP disseminates these reports throughout VA and in the published literature; some evidence syntheses have informed the clinical guidelines of large professional organizations.

The ESP Coordinating Center (ESP CC), located in Portland, Oregon, was created in 2009 to expand the capacity of QUERI/HSR&D and is charged with oversight of national ESP program operations, program development and evaluation, and dissemination efforts. The ESP CC establishes standard operating procedures for the production of evidence synthesis reports; facilitates a national topic nomination, prioritization, and selection process; manages the research portfolio of each Center; facilitates editorial review processes; ensures methodological consistency and quality of products; produces “rapid response evidence briefs” at the request of VHA senior leadership; collaborates with HSR&D Center for Information Dissemination and Education Resources (CIDER) to develop a national dissemination strategy for all ESP products; and interfaces with stakeholders to effectively engage the program.

Comments on this evidence report are welcome and can be sent to Nicole Floyd, ESP CC Program Manager, at Nicole.Floyd@va.gov.

Recommended citation: Rector T, Taylor BC, Sultan S, Shaukat A, Adabag S, Nelson D, Capecchi T, MacDonald R, Greer, N, Wilt TJ. Life Expectancy Calculators, VA ESP Project #09-009; 2016.

This report is based on research conducted by the Evidence-based Synthesis Program (ESP) Center located at the Minneapolis VA Medical Center, Minneapolis, MN, funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Quality Enhancement Research Initiative. The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. No investigators have any affiliations or financial involvement (*eg*, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

EXECUTIVE SUMMARY

INTRODUCTION

Estimation of life expectancy serves an important role in clinical decisions about screening for disease and treatment in primary care practices. Increasingly, clinical guidelines recommend a course of action in context of the patient's life expectancy. While many tools have been developed to estimate life expectancy or the likelihood of surviving for a period of time, there are no well-established decision support tools that are broadly applicable to primary care.

For a life expectancy calculator to gain widespread use in primary care, it needs to have acceptable and validated predictive accuracy at decision-relevant times such as 6 months or 5 years, proven usefulness, and be readily available. In 2012, a systematic review by Yourman et al summarized existing life expectancy calculators that may be useful for older patients in primary care. However, none of the reviewed calculators met our criteria for widespread use. The purpose of this systematic review was to update the search on life expectancy calculators, with focus on calculators that may be valuable for use in the Veterans Health Administration population receiving primary care.

With input from Operational Partners and Technical Expert Panel members, we developed the following Key Questions for the review:

Key Question 1: Between 2011 and 2016, have there been any additional reports of life expectancy calculators that may have sufficient predictive accuracy for use in adult primary care practice?

Key Question 2: Of the life expectancy calculators being reviewed have any external validation studies been published between 2011 and 2016? If yes, what population was studied and what was the predictive accuracy therein?

Key Question 3: What is the clinical use of the mortality prediction models (*aka* life expectancy calculators), and was there improvement in patient survival times, health-related quality of life, provider-patient communication, patient satisfaction and participation in clinical decisions, or lower healthcare utilization and costs?

METHODS

Data Sources and Searches

We searched MEDLINE (Ovid) from 2011 to May 2016 using title words for life expectancy, calculators or models, survival, mortality, death, and validation or calibration. The search was limited to English language and studies of humans middle-aged (45 plus years) and older. We also limited the search to relevant study designs.

Study Selection

The conceptual basis for this review was that a decision support tool would be electronically implemented to provide quantitative survival estimates to primary care providers who along with their patients would use the estimates when making healthcare decisions and thereby improve patient outcomes and healthcare value. We included studies of methods to estimate probabilities of survival or dying within a specified period of time. In addition, we were interested in studies of the use and effects of the resulting estimates.

We required that studies offer a method of estimation based on variables that would be generally available for primary care outpatients. Thus studies of estimates that require inpatient variables or patient surveys, studies of novel biomarkers, or studies of individual predictors, including those based on only age, sex, and race, were excluded. Studies that were limited to tests for associations such as relative risks, odds ratios, and hazard ratios were considered preliminary work and excluded.

Data Abstraction, Quality Assessment, Synthesis, and Strength of Evidence

Information on study sample and prediction model characteristics and performance were abstracted into evidence tables by one trained investigator and verified by the principal investigator.

We assessed quality of each study using selected items from the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS). Five criteria were selected: 1) predictor definition/measurement, 2) outcome definition/measurement, 3) independence of outcome and predictor assessments, 4) completeness of follow-up/predictor data, and 5) validation.

The evidence is narratively described without any formal meta-analysis. The strength of evidence for each key question is rated as high, low, or insufficient based on the number of quality studies and the consistency of the results.

RESULTS

Our literature search yielded 8,120 potentially eligible studies. Titles were reviewed by trained investigators and 509 studies were selected for abstract review. Each abstract was reviewed by 2 trained investigators and 51 studies were identified for full-text review. Full-text articles were reviewed by an investigator and the principal investigator. We included 10 studies from the literature search and an additional study identified from hand-searching studies included in the full-text review phase.

Key Question 1

We found 8 different mortality prediction models (in 11 reports). All had encouraging predictive accuracy. All but one model used patient's age and sex along with diagnostic codes extracted from electronic databases with or without additional variables. Models developed for mortality risk adjustment focused on mortality during the first year of follow-up while models intended for decisions about cancer screening had longer time horizons of 5 to 10 years. The best-characterized model based on VHA data, known as the Care Assessment Needs (CAN) score,

was not strictly a life-expectancy model as it focused on deaths without a preceding hospitalization rather than all deaths.

Overall, the 8 models were reasonably well-calibrated, reporting relatively small differences between observed and predicted mortality in arbitrary mortality risk groups. Thus, if a clinician used one of the models to place a patient in a mortality risk group, the average or median predicted mortality for the risk group could be used as an estimate. However, some individual predicted probabilities of survival may be substantially different from the average of the risk group. Furthermore, the differences in estimates between risk groups may not be sufficient to prompt clinicians or patients to make different decisions about screening or treatment.

The strength of evidence of the body of evidence for Key Question 1 was rated as high based on finding several studies of acceptable quality that reported a mortality prediction model based on similar predictor variables that had potentially useful predictive accuracy.

Key Question 2

One of the life expectancy models identified in Key Question 1 was in a sense validated in the older patients from the population used to generate the model and a sub-population of individuals with schizophrenia. Another model was in a sense validated in the older patients and a sub-population of heart failure patients. Calibration was not always as good in the sub-populations.

Not finding any true external validation studies of the prediction models in primary care practices, we rated the strength of evidence for Key Question 2 as insufficient.

Key Question 3

We found no studies that examined the effects of using the reviewed mortality prediction models in clinical care.

The strength of evidence for Key Question 3 was rated as insufficient.

DISCUSSION

Key Findings

- Between 2011 and 2016, 11 studies reported on 8 mortality prediction models; all but one of the included studies (a study from Japan) utilized data from large electronic databases.

- Models were developed for different purposes including development of individual risk scores to adjust for possible difference in mortality risk when comparing healthcare outcomes, to help primary care teams assesses short-term risk of hospitalization or death without hospitalization, or to help healthcare providers judge whether patient will or won’t die within a specified time pertinent to decisions about screening for cancer.

- C-statistics ranged from 0.77 to 0.90, indicating the models provide good discrimination of those who survived or died during the varying periods of follow-up. Few studies reported the sensitivity, specificity, or positive or negative predictive values for a proposed risk score threshold that would be used to determine which patients should or shouldn’t be screened or treated.

- The prediction models were generally well-calibrated to the test samples with seemingly insubstantial differences between observed and predicted mortality across a range of risk groups.

- We found no true external validation studies of the reviewed mortality prediction models. None of the models have been externally validated for general primary care use.

- No studies meeting eligibility criteria for the review examined the impact of using one of the life expectancy calculators to improve healthcare decisions or outcomes.

Applicability

Several of the prediction models identified for this review could potentially be adapted for use in VA primary care practices. However, the discrimination and calibration of the models may vary in VA primary care practices that may have substantially different distributions of the predictors and/or mortality, and therefore would need to be evaluated. To adapt a mortality prediction model, the sources of data and definitions of predictor variables will need to be consistent across VA healthcare systems and over time. Centralized prediction models that are made widely available should be calibrated to and validated in several VA primary care patient populations and periodically checked and recalibrated if necessary. Although feasible, we found no studies to indicate whether making a reasonably accurate life expectancy calculator available to VA primary care providers or patients would influence their healthcare decisions or outcomes.

Research Gaps/Future Research

Although healthcare providers and guidelines make recommendations based, in part, on assessments of life expectancy, there is no widely accepted statistical tool for estimating patients' life expectancy or probability of survival, particularly for prolonged periods, for example, 10 years. Research on the clinical usefulness and impact of mortality prediction models on clinical decisions and outcomes is needed to guide further development and engender widespread acceptance. Analytical life expectancy predictions have to be demonstrated to be more accurate than clinicians' intuitive prognostic assessments, preferably using statistics that allow clinicians to compare prognostic errors. Additionally, they should use patient information that is readily and reliably available. Strong comparative evidence that using a quantitative prediction model can reduce practice variation and improve healthcare decisions and outcomes most likely will be needed to change current practices.

Conclusions

Life expectancy calculators based on readily available electronic data that have acceptable performance for estimating one-, 5-, and 10-year life expectancy in middle age to older adults are feasible. These calculators need to be validated for use in primary care practices. There are no data on the effect of using these life expectancy calculators on patient or provider decisions or outcomes. If a life expectancy calculator is made available, it remains to be determined whether primary care providers would use it or whether it would improve healthcare delivery, resource use, patient experiences, or outcomes.