

Life Expectancy Predictions for Older Diabetic Patients as Estimated by Physicians and a Prognostic Model

Claire E. O'Hanlon, MPP, Jennifer M. Cooper, MPH, Sang Mee Lee, PhD, Priya John, MPH, Matthew Churpek, MD, MPH, Marshall H. Chin, MD, MPH, Elbert S. Huang, MD, MPH

Background: Multiple medical organizations recommend using life expectancy (LE) to individualize diabetes care goals. We compare the performance of patient LE predictions made by physicians to LE predictions from a simulation model (the Chicago model) in a cohort of older diabetic patients. **Design:** Retrospective cohort study of a convenience sample ($n = 447$) of diabetes patients over 65 years and their physicians. **Measurements:** Physicians provided LE estimates for individual patients during a baseline survey (2000–2003). The prognostic model included a comprehensive geriatric type 2 diabetes simulation model (the Chicago model) and combinations of the physician estimate and the Chicago model (“And,” “Or,” and “Average” models). Observed survival was determined based on the National Death Index through 31 December 2010. The predictive accuracy of LE predictions was assessed using c -statistic for 5-year mortality; Harrell’s c -statistic, and Integrated Brier score for overall survival. **Results:** The patient cohort had a

mean (SD) age of 73.4 (5.9) years. The majority were female (62.6%) and black (79.4%). At 5 years, 108 (24.2%) patients had died. The c -statistic for 5-year mortality was similar for physicians (0.69) and the Chicago model (0.68), while the average of estimates by physicians and Chicago model yielded the highest c -statistic of any method tested (0.73). The estimates of overall survival yielded a similar pattern of results. **Limitations:** Generalizability of patient cohort and lack of updated model parameters. **Conclusions:** Compared with individual methods, the average of LE estimates by physicians and the Chicago model had the best predictive performance. Prognostic models, such as the Chicago model, may complement and support physicians’ intuitions as they consider treatment decisions and goals for older patients with chronic conditions like diabetes. **Key words:** diabetes, aged, life expectancy, prediction, computer simulation, physicians. (MDM Policy & Practice XXXX;XX:1–9)

The growing population of older people with diabetes presents a significant public health problem around the world.¹ In the United States, the number of people over 65 diagnosed with diabetes is expected to rise from 8.1 million patients in 2000 to 16.8 million by 2050.² Diabetes treatment goals have historically aimed to achieve near-normal levels of glucose (hemoglobin A1c [HbA_{1c}] <7%), blood pressure (<130/80 mmHg), and cholesterol to reduce the risk of complications. However, these goals are inconsistently achieved in practice³ and often require intensive treatment.⁴ Clinical trials that inform diabetes recommendations^{5–7} usually

exclude older and sicker patients.⁸ Thus, it is unclear what glycemic target goal clinicians should aim for in the older patient, as this population is highly heterogeneous in terms of complications, comorbidities, and functional status,⁹ and these factors influence a patient’s health goals and the potential outcomes of intensive treatment. Patient-centered care for older patients with diabetes therefore requires individualized glycemic targets.

Recognizing the need to individualize diabetes care, multiple medical organizations have recommended using life expectancy (LE) as an approach to selecting optimal glycemic targets for elderly patients. In 2003, an American Geriatrics Society/California Healthcare Foundation panel issued one of the earliest guidelines recommending the use of LE to guide individualization of diabetes care in older patients.¹⁰ They recommended that patients

© The Author(s) 2017

Reprints and permission:

<http://www.sagepub.com/journalsPermissions.nav>

DOI: 10.1177/2381468317713718



This Creative Commons Non Commercial CC-BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

with limited LE (5 years or less) to live should strive for moderate glucose control ($HbA_{1C} < 8\%$), while patients with greater LE should pursue more intensive goals, as observable benefits require 9 years of ongoing intensive glucose control.⁵ In 2012, the European Association for the Study of Diabetes and the American Diabetes Association both recommended an individualized approach to selecting glucose control targets based on variables such as decreasing LE, increasing cognitive and functional impairment, and frailty.^{11,12}

While multiple guidelines now advise using LE in making medical decisions, carrying out this recommendation in practice is both ethically controversial and technically challenging. Physicians and patients can provide LE estimates, but prior studies on LE estimation have often found that both physicians and patients are generally overly optimistic.^{13–19} In an effort to systematically generate LE predictions, we created a simulation model for older patients with type 2 diabetes, the “Chicago model,” that produces estimates of patient LE and the risk of complications. While a number of mortality prediction models are available,²⁰ the present study aimed to externally validate the prognostic accuracy of the Chicago model compared with the status quo of prognostication: the physician’s judgment. We compare the accuracy of predictions of

patients’ LE (5-year mortality and overall survival) estimated by the Chicago model, the patient’s physician, and several combinations of these estimates in a cohort of older patients with diabetes.

METHODS

Patient Cohort

Patients with diabetes age 65 and older were enrolled in a study of treatment preferences and goals between December 2000 and January 2003 as previously described.²¹ Eligible patients were identified before visiting a University of Chicago internal medicine, geriatrics, or endocrinology clinic. Of 1,067 potential participants telephoned, 694 answered and 607 agreed to participate. Fifty-two patients were no-shows, leaving 555 who were surveyed (80% of approached patients). Of these, 108 were excluded from this analysis if their medical and/or death record was unverifiable (8 patients), LE predictions were missing (94 patients), or they fulfilled multiple exclusion criteria (6 patients). Four hundred forty-seven patients constituted the final patient cohort. Patients’ clinical values, if available, were extracted from the medical records and used for risk factors and biomarkers in the Chicago model. Average values from the National Health and Nutrition Examination Survey²² based on patient age, sex, and race were used to replace missing data. This study was approved by the University of Chicago Institutional Review Board.

The Chicago Model

The Chicago Type 2 Geriatric Diabetes Simulation Model (“Chicago model”) is a Markov Chain Monte Carlo simulation model created using Microsoft Excel-based @Risk 4.5 (Palisade Corporation, Newfield, NY). The model’s general structure has been previously described.²³ This model integrates the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model,²⁴ a series of integrated diabetes complication models based on the UKPDS cohort, and a 4-year mortality index developed from the Health and Retirement Study.²⁵ The Chicago model accounts for demographics, functional status, comorbid illness, risk factors, and duration of diabetes. The initial LE for each patient was obtained by running the Chicago model for 1,000 repetitions, and odds of death were multiplied by 2.75 to account for higher background

Received 17 February 2017 from the Pardee RAND Graduate School, Santa Monica, California (CEO); Section of General Internal Medicine (CEO, JMC, PJ, MHC, ESH), Section of Pulmonary Critical Care (MC), and Department of Public Health Sciences (SML), University of Chicago, Chicago, Illinois; Department of Obstetrics and Gynecology, Northwestern University Feinberg School of Medicine, Chicago, Illinois (JMC); San Francisco Health Plan, San Francisco, California (PJ). This article was presented in part at the 2013 Society of General Internal Medicine Annual Meeting in Denver, Colorado, and the Mount Hood 2014 Challenge in Stanford, California. This study was funded by a pilot grant from the National Opinion Research Center (NORC) Center on Demography and Economics of Aging at the University of Chicago (P30 AG012857), the National Institute of Diabetes and Digestive and Kidney Diseases Chicago Center for Diabetes Translation Research (P30 DK092949), and the National Institute on Aging (R01AG17872). MHC and ESB were also supported by a National Institute of Diabetes and Digestive and Kidney Diseases Midcareer Investigator Awards in Patient-Oriented Research (K24 DK071933, K24 DK105340). The funding agreement ensured the authors’ independence in designing the study, interpreting the data, writing, and publishing the report. Revision accepted for publication 17 April 2017.

Address correspondence to Elbert S. Huang, MD, MPH, The University of Chicago, 5841 S. Maryland Avenue, MC 2007, Chicago, IL 60637, USA; telephone: 773-834-9143; e-mail: ehuang@medicine.bsd.uchicago.edu.

mortality rates in patients with diabetes.²⁶ The estimated LE was transformed into a binary indicator of limited LE (≤ 5 years) if applicable for the analysis.

It is important to note that data on three functional status measures that are integrated into the 4-year mortality index were not collected in the patient survey. Two questions (“Because of a health or memory problem, do you have any difficulty with managing your money—such as paying your bills and keeping track of expenses?” and “Because of a health or memory problem, do you have any difficulty with pulling or pushing large objects like a living room chair?”) were approximated with other survey questions (Mini-Mental Status Examination²⁷ score < 17 or diagnosis of dementia and “How much does your health now limit you in doing moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf?”). One question (“Because of a health or memory problem, do you have any difficulty with walking several blocks?”) was omitted because no analogous question existed.

Physician Estimates

For each patient enrolled, a one-page survey was distributed to the patient’s physician after the visit. The surveyed physician had to be responsible for the patient’s diabetes care and willing to complete the survey for each patient enrolled in the study. Seventy-seven of 79 eligible physicians agreed to complete the survey. After patient exclusions, 63 physicians (10 geriatricians, 38 internists, and 15 endocrinologists) provided estimates. Physicians filled in a blank in response to the question, “How many years do you estimate that this patient will live?” Similar to the Chicago model’s estimate, this value was either made into a binary indicator for limited LE (≤ 5 years) or kept as a continuous value.

Combinations of Physician Estimates and Chicago Model: The “And,” “Or,” and “Average” Models

We considered three additional methods of generating binary classifications of “limited LE” (≤ 5 years) to approximate how a physician might interact with a prognostic model in practice to decide whether or not a patient’s LE should be considered limited. In these models, estimates by the model and physician are combined according to several different decision rules. These models are referred to as the “And” model (both the physician and

model predict limited LE), the “Or” model (either the physician or model predict limited LE), and the “Average” model (limited LE is determined by the mean of the physician and model’s estimates). In the “And” model, a patient is considered to have limited LE if and only if both the physician and the Chicago model generate LE predictions that are less than or equal to 5 years. In the “Or” model, a patient is considered to have limited LE if either the physician or the Chicago model generate an LE prediction that is less than or equal to 5 years. Finally, in the “Average model,” the mean of LE estimates from the physician and the Chicago model is generated. If the mean of these estimates is less than or equal to 5 years, the patient is considered to have limited LE. The “Average model” also provides a third point estimate of the patient’s LE, in addition to the binary classification of limited LE. These categorization schemes are based on how we believe physicians interact with the model or other prognostic information in a real-life clinical setting.

Observed Survival

Observed survival was determined by querying the National Death Index (NDI) from 1 January 2001 to 31 December 2010. Patients’ first and last names, birth date, last known state of residency, and last date of contact were submitted to the NDI. All information submitted matched NDI records for 172 subjects (“perfect matches”). Twenty-eight submissions returned as “partial matches” due to missing or incorrect information were judged to be matches after comparison to electronic medical records. One additional date of death was identified through electronic medical records.

Statistical Analysis

When assessing the prognostic accuracy of physician and model estimates, we used several different measures depending on the type of predictive estimate; for 5-year mortality (a binary indicator for death at or before 5 years from the date of survey), we calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), receiver-operating characteristic (ROC) curve, and *c*-statistic. The *c*-statistic is equivalent to the area under the receiver-operating characteristic curve (AUC), incorporating both sensitivity and specificity (a *c*-statistic of 0.5 indicates predictive

Table 1 Patient Characteristics (*N* = 447)

<i>Sociodemographic characteristics</i>	
Female, <i>n</i> (%)	280 (62.6)
Race, <i>n</i> (%) ^a	
Black	355 (79.4)
White	66 (14.8)
Asian	2 (0.5)
Hispanic	13 (2.9)
Other	11 (2.5)
Age (years), mean (SD)	73.4 (5.9)
Age, <i>n</i> (%) ^a	
65–69	137 (30.7)
70–74	125 (28.0)
75–79	115 (25.7)
80–94	70 (15.7)
Education, <i>n</i> (%) ^a	
Less than high school	131 (29.3)
High school graduate	115 (25.7)
High school and higher	198 (44.3)
No answer	3 (0.7)
Self-reported income, <i>n</i> (%) ^a	
<\$10,000	102 (22.9)
\$10,001–\$25,000	118 (26.5)
\$25,001–\$50,000	102 (22.9)
>\$50,000	53 (11.9)
Did not know/refused	71 (15.9)
Marital status, <i>n</i> (%) ^a	
Married or living as if married	194 (43.4)
Divorced/separated/widowed	236 (52.8)
Single (never married)	17 (3.8)
<i>Clinical characteristics</i>	
Type 2 diabetes, <i>n</i> (%)	446 (99.8)
Duration of diabetes mellitus (years), mean (SD)	13.2 (10.4)
Duration of diabetes mellitus (years), <i>n</i> (%) ^a	
0–5	132 (29.5)
6–10	96 (21.5)
11–15	80 (17.9)
15–20	47 (10.5)
>20	92 (20.6)
Body mass index, mean (SD)	30.0 (6.5)
Hemoglobin A1c (%), mean (SD)	7.65 (1.63)
Hemoglobin A1c <7%, <i>n</i> (%), <i>N</i> = 439	157 (35.7)
Systolic blood pressure (mmHg), mean (SD)	143.6 (22.6)
Systolic BP <130 mmHg, <i>n</i> (%), <i>N</i> = 434	101 (23.3)
LDL cholesterol (mg/dL), mean (SD)	106.1 (39.5)
LDL cholesterol <100 mg/dL, <i>n</i> (%), <i>N</i> = 336	169 (50.3)
Total number of diabetes medicines, mean (SD)	0.9 (0.8)
Total number of hypertension medicines, mean (SD)	1.6 (0.9)
Total number of lipid-lowering medicines, mean (SD)	0.4 (0.5)

*(continued)***Table 1** (continued)

Complications, <i>n</i> (%)	197 (44.1)
Nephropathy	80 (17.9)
Neuropathy	93 (20.8)
Retinopathy	57 (12.8)
Peripheral vascular disease	83 (18.6)
Charlson Comorbidity Index score, mean (SD)	3.07 (1.81)
<i>Mortality outcomes</i>	
Deceased by the end of 5 years, <i>n</i> (%)	108 (24.2)
Deceased by the end of the study period, <i>n</i> (%)	201 (45.0)

a. Percentages may not add up to 100 due to rounding.

value no better than random chance and *c*-statistic of 1.0 indicates perfect prediction). To evaluate the accuracy of LE estimates overall (survival time), we first censored those who lived at the end of observation period. The observed survival time was compared with predicted estimates from physician and four models using time-dependent ROC curve,²⁸ Harrell's *c*-statistic,²⁹ and integrated Brier score (IBS).^{30,31} Harrell's *c*-statistic is an adapted *c*-statistic that accounts for length of survival. The integrated Brier score is an overall measure for the inaccuracy of a prediction model whose value ranges from 0 to 1 (smaller values indicate better performance and a value of 0.25 indicates a lack of discriminatory power). The measure is a weighted average of the squared distance between the observed survival status and the predicted survival probability of a model. Among different evaluation metrics, *c*-statistic, the most commonly used, was considered as the primary method in this article. All analyses were conducted using R version 3.0.1.

RESULTS

Baseline Characteristics and Properties of LE Estimates

The mean (SD) follow-up period was 8.9 (0.6) years (range = 7.9–10.1 years) for the 447 patients in the study cohort (Table 1). One hundred eight (24.2%) subjects died at or before 5 years from the date of survey, while 201 (45%) subjects died by the end of the follow-up period. Properties of LE estimates differed by source (Figure 1). Physicians predicted the highest average LE of 9.3 (4.8) years, the widest range (0.5–30 years), while the Chicago model had the shortest mean LE prediction (7.3 (3.7) years, range = 1.1–18.6 years). The mean LE

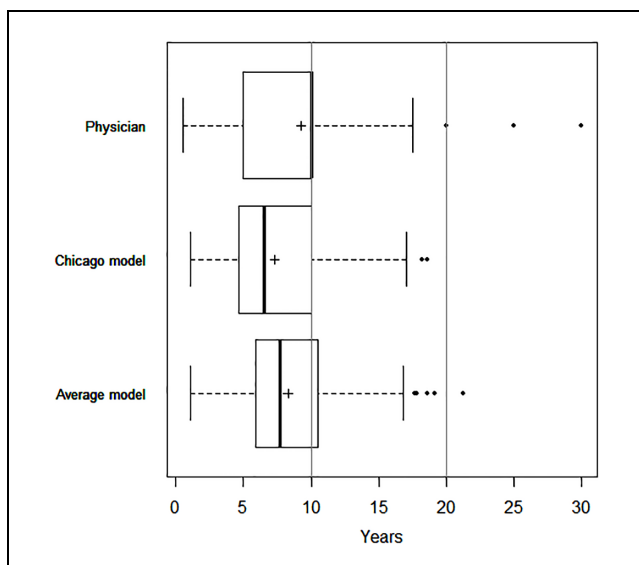


Figure 1 Boxplot display of life expectancy (LE) predictions by patients, physicians, and prognostic models. Boxplot displays the median (bold vertical line), interquartile range (IQR; solid line box), 1.5 IQR adjacent values (whiskers), outliers (points), and the mean value (+) for LE predictions by physicians (“How many years do you estimate that this patient will live?”), the Chicago model, and the Average model (mean of physician’s LE prediction and LE output of Chicago model).

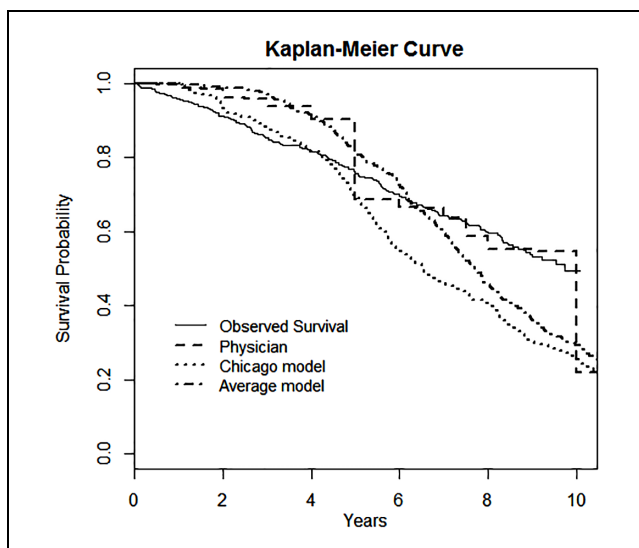


Figure 2 Survival outcomes of observed patient death versus predicted survival by physician, the Chicago model, and the average model.

estimate of the combination of the physician and model’s estimates was 8.3 (3.5) years (range = 1.0–21.3 years). Physicians tend to predict LE in integer

Table 2 Comparison of Life Expectancy (LE) Predictions by Physician, Chicago Model, and Observed Patient Death (N = 447)

	Observed Death	
	Alive at 5 Years, n (%)	Dead at 5 Years, n (%)
Physician’s prediction		
Alive at 5 years	256 (57)	51 (11)
Dead at 5 years	83 (19)	57 (13)
Chicago model’s prediction		
Alive at 5 years	253 (57)	57 (13)
Dead at 5 years	86 (19)	51 (11)

values and their estimates cluster in 5-year increments (e.g., 5, 10, 15 years, etc.), while the model outputs more evenly distributed across the range (see Figure 2 for predicted and observed survival curves).

The physicians and the Chicago model were not divergent in their overall estimates of 5-year mortality, with the physician guessing correctly 70% of the time to the Chicago model’s 68% (Table 2). For approximately two thirds (66%) of study subjects, the prediction of 5-year mortality generated by the physician and the model were the same ($n = 62$ for 5-year mortality and $n = 232$ for $LE > 5$ years), and differed in about a third (34%) of subjects ($n = 75$ for only the model predicting 5-year mortality and $n = 78$ for only the physician predicting 5-year mortality). Among 75 study subjects for whom the model predicted 5-year mortality and the physician did not, 31% were deceased at 5 years and 69% were alive. Among 78 study subjects for whom the physician predicted 5-year mortality and the model did not, 37% were deceased at 5 years and 63% were alive.

Predictive Performance for 5-Year Mortality

Properties of predictions by the Chicago model, physicians, and combinations of the Chicago model and physician estimates (“And,” “Or,” “Average” models) of 5-year mortality (death by 5 years from date of survey) were calculated (Table 3). The physician and Chicago model estimates performed similarly, with the performance of the “Average” model exceeding the Chicago model and physician estimates, as well as the “And” and “Or” models for 5-year mortality according to the *c*-statistic, though these estimates

Table 3 Performance Metrics of Life Expectancy (LE) Predictions for 5-Year Mortality and Overall Survival Time by Physicians, the Chicago Model, the Average of the Physician and the Chicago Model, the “And” Model, and the “Or” Model

	5-Year Mortality					Overall Survival	
	Sensitivity	Specificity	PPV ^a	NPV ^b	<i>c</i> -Statistic ± SE (95% CI) ^c	Harrell's <i>c</i> -Statistic ± SE (95% CI) ^d	IBS ^e
Physician ^f	0.53	0.76	0.41	0.83	0.69 ± 0.027 (0.67–0.74)	0.65 ± 0.027 (0.64–0.74)	0.148
Chicago model ^g	0.47	0.75	0.37	0.82	0.68 ± 0.028 (0.63–0.78)	0.66 ± 0.028 (0.63–0.74)	0.147
Average model ^h	0.35	0.87	0.46	0.81	0.73 ± 0.026 (0.68–0.78)	0.69 ± 0.026 (0.68–0.78)	0.140
And model ⁱ	0.26	0.90	0.45	0.79	0.58 ± 0.023 (0.53–0.63)	NA ^j	NA ^j
Or model ^k	0.74	0.60	0.37	0.88	0.67 ± 0.024 (0.62–0.72)	NA ^j	NA ^j

Note: PPV = positive predictive value; NPV = negative predictive value; SE = standard error; CI = confidence interval; IBS = integrated Brier score.

a. Positive predictive value: true positives/(true positives + false positives).

b. Negative predictive value: true negatives/(true negatives + false negatives).

c. Area under the receiver operating characteristic (ROC) curve (AUC or *c*-statistic). The *c*-statistic is equivalent to the area under the receiver-operating characteristic curve, incorporating both sensitivity and specificity (a *c*-statistic of 0.5 indicates predictive value no better than random chance and *c*-statistic of 1.0 indicates perfect prediction).

d. Harrell's *c*-statistic.²⁹

e. Integrated Brier score, a predictive accuracy score function that takes one values between 0 and 1, with lower values indicating better predictive performance.^{30,31}

f. Physician's answer on a patient-specific questionnaire to the question, “How many years do you estimate that this patient will live?” The physician's answer to this question was converted to a binary indicator of predicted five-year mortality when the answer to this question was 5 years or less.

g. Chicago model generates a point estimate of the patient's LE which is turned into a binary indicator of 5-year mortality when it was equal to 5 years or less.

h. Average model takes the mean of the point estimates generated by the physician and the Chicago model and uses this as its predictor. The average model is also converted into a binary indicator of predicted 5-year mortality when the average is equal to 5 years or less.

i. “And” model predicts 5-year mortality when both the physician and the Chicago model predict 5-year mortality.

j. NA = not applicable. Performance metrics for overall survival time cannot be computed for these models because they generate a binary classification of limited LE rather than a point estimate of LE.

k. “Or” model predicts 5-year mortality when either the physician or the Chicago model predict 5-year mortality.

had wide confidence intervals and differences were not statistically significant. As expected by the way these models are defined, the “Or” model had the highest sensitivity (0.74) and NPV (0.88) but the lowest specificity (0.60), while the “And” model had the lowest sensitivity (0.26) and NPV (0.79) but the highest specificity (0.90). The “And” model had the lowest (worst) *c*-statistic for 5-year mortality (0.58), followed by the “Or” model (0.67) and the Chicago model (0.68). The *c*-statistic generated for the physician estimate (0.69) was slightly higher. The “Average” model outperformed all of the other models, generating the highest *c*-statistic (0.73) for 5-year mortality.

Predictive Performance for Overall Survival

Metrics to compare performance of physicians, the Chicago model, and the average of the Chicago model estimate and physician estimate for predicting overall survival time (Harrell's *c*-statistic and IBS) were also calculated (Table 3). See Figure 3A and B for time-dependent ROC curves for overall survival. The Harrell's *c*-statistic was similar for the physician (0.65), the Chicago model (0.66), and the average of the Chicago model and physician (0.69). Similarly, physician estimates had the highest

(worst) IBS (0.148); the mean of the physician and model estimate had the lowest (best) (0.140).

DISCUSSION

The inaccuracies of physicians and patients in predicting LE are well-established, and prognostic models are a promising approach, with numerous models and calculators having been developed (<http://eprognosis.ucsf.edu/>).²⁰ An important limitation of these models is that they are always developed using data from past observations, and therefore are the results of the historical timing of diagnosis, natural history of disease, and effects of past treatments, all of which evolve over time. Given the limitations of both physicians and these models, it is critical to examine how their predictions may actually be used in real clinical situations. Our study provides insight into how a simulation model of LE in older patients with diabetes performs in a specific patient population and how this model's predictions compare to, interact with, and supplement the judgment of physicians.

Predictions of LE made by physicians are admittedly subjective, and we did not systematically examine whether a physician's experience or

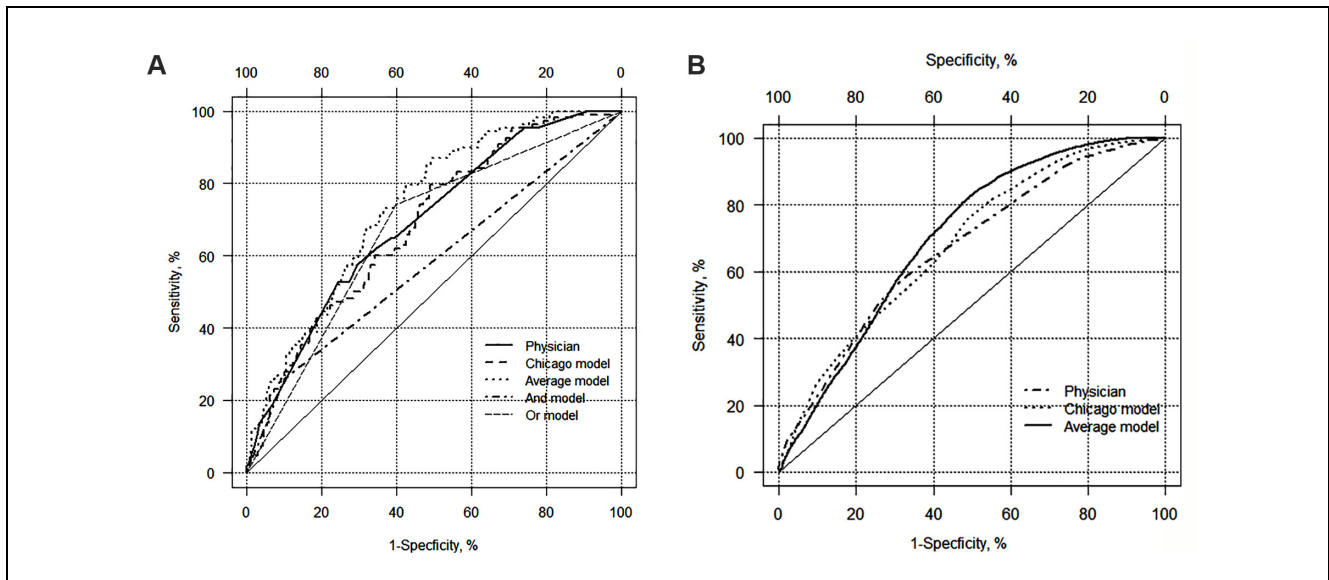


Figure 3 (A) ROC curve for 5-year mortality. The diagonal solid line indicates a test with no discriminatory power (area under the curve equal to 0.5). (B) ROC curve for overall survival time. The diagonal solid line indicates a test with no discriminatory power (area under the curve equal to 0.5).

expertise was correlated with prognostic accuracy in this study. However, physicians have access to other information that may be critical for predicting LE, such as a patient's resilience or social support, that are harder to operationalize as model inputs. When we compared their estimates to the Chicago model, which is a relatively complex microsimulation model that accounts for multiple clinical domains as well as comorbidity and functional impairment, we did not find substantial differences in predictive performance for 5-year mortality or survival duration. Both methods of LE prediction had relatively low sensitivity, somewhat higher specificity, and had similar *c*-statistics for predicting 5-year mortality. While the different combination models we evaluated varied in their predictive performance, they had modest predictive power at best, and point toward deficiencies in what we know about prognosis. However, the improved performance of the "Or" model with respect to sensitivity and the average model with respect to the *c*-statistic suggest that physician and model estimates may be best used in concert rather than independently.

At present, we suspect that physicians do not frequently differentiate diabetes goals and treatments based on prognosis. We have previously found, in our own practice, that glycemic control levels are nearly identical for older patients with different

health status based on physician observations and that the sickest patients are on the most intensive regimens.³² The tendency of physicians to not alter their actions in response to relevant information, or status quo bias,³³ may lead to overtreatment of patients unlikely to benefit from intensive therapy or at high risk for adverse events,^{34–36} as well as undertreatment of patients likely to benefit from intensive therapy. If physicians were consistently provided with prognostic information at the point of care, they might be able to overcome status quo bias and make clinical decisions more tailored and in tune with the health status and trajectory of individual patients. Within the context of this study, in the two thirds of patients for whom the physician and Chicago model predictions were concordant, the model could have helped confirm the physician's intuition and treatment plan. However, for the one third of patients for whom predictions were discordant, the physician faces a dilemma of how to use the new information provided by the model—ignore it, adopt it, or somehow incorporate it. In this study, both types of discordance resulted in about a one third chance of 5-year mortality. If a physician's prognostication can be used in tandem with external prediction, it may lead to increased accuracy of LE prediction and increase willingness to consider prognostic models.

This study utilizes a unique data set of LE estimates from physicians made on an individual-patient basis to externally validate the predictive performance of a simulation model compared to the physician's "off-the-cuff" estimate. Limitations of this study include that subjects received care at a single academic medical institution (University of Chicago Medicine) and that the prognostic abilities of physicians may not be representative of physicians elsewhere. Patients are also largely drawn from a specific demographic and socioeconomic group (mostly African Americans residing in Chicago's South Side) distinct from the populations used to create the model; the same exercise in another population might look different. Last, the Chicago model does not yet incorporate updated estimates of diabetes complications and mortality,³⁷ which may better reflect current diabetes care practices.

Despite the proliferation of new recommendations to incorporate LE into medical decisions, including whether to prescribe medications,³⁸ undergo cancer screenings,³⁹ and refer patients to hospice care,¹³ the controversies and challenges of estimating prognosis in clinical practice have been infrequently studied. Predicting the LE of patients with acute conditions and short survival horizons remains difficult, and LE prediction may be even more difficult in older patients with chronic conditions like diabetes, in which long survival horizons are possible.⁴⁰ Advances in health information technology and the increased presence of computers in primary care settings have led to experimentation with real-time decision support tools integrated into electronic medical records.⁴¹ In the near future, LE estimates from prognostic models could be similarly integrated and referenced during clinical encounters. The modest performance of the Chicago model provides evidence for the need to develop and validate a more contemporary model of diabetes complications that reflects the diversity of the United States. Even with the limitations of current data sources, our results suggest that prognostic models have the potential to complement and support physicians as they work with patients to make complex decisions about chronic disease management at advanced ages.

REFERENCES

1. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care*. 2004;27(5):1047–53.
2. Boyle JP, Honeycutt AA, Narayan KM, et al. Projection of diabetes burden through 2050: impact of changing demography and disease prevalence in the U.S. *Diabetes Care*. 2001;24(11):1936–40.
3. Saaddine JB, Cadwell B, Gregg EW, et al. Improvements in diabetes processes of care and intermediate outcomes: United States, 1988–2002. *Ann Intern Med*. 2006;144(7):465–74.
4. Grant RW, Pirraglia PA, Meigs JB, Singer DE. Trends in complexity of diabetes care in the United States from 1991 to 2000. *Arch Intern Med*. 2004;164(10):1134–9.
5. U.K. Prospective Diabetes Study Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet*. 1998;352(9131):837–53.
6. U.K. Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. *BMJ*. 1998;317:703–13.
7. Gaede P, Lund-Andersen H, Parving HH, Pedersen O. Effect of a multifactorial intervention on mortality in type 2 diabetes. *N Engl J Med*. 2008;358(6):580–91.
8. Gerstein HC, Miller ME, Byington RP, et al. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med*. 2008;358(24):2545–59.
9. Blaum CS. Management of diabetes mellitus in older adults: are national guidelines appropriate? *J Am Geriatr Soc*. 2002;50(3):581–3.
10. Brown AF, Mangione CM, Saliba D, Sarkisian CA, California Healthcare Foundation/American Geriatrics Society Panel on Improving Care for Elders with Diabetes. Guidelines for improving the care of the older person with diabetes mellitus. *J Am Geriatr Soc*. 2003;51(5 Suppl Guidelines):S265–80.
11. Kirkman MS, Briscoe VJ, Clark N, et al. Diabetes in older adults. *Diabetes Care*. 2012;35(12):2650–64.
12. Sinclair A, Morley JE, Rodriguez-Manas L, et al. Diabetes mellitus in older people: position statement on behalf of the International Association of Gerontology and Geriatrics (IAGG), the European Diabetes Working Party for Older People (EDWPOP), and the International Task Force of Experts in Diabetes. *J Am Med Dir Assoc*. 2012;13(6):497–502.
13. Christakis NA, Lamont EB. Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ*. 2000;320(7233):469–72.
14. Vigano A, Dorgan M, Bruera E, Suarez-Almazor ME. The relative accuracy of the clinical estimation of the duration of life for patients with end of life cancer. *Cancer*. 1999;86(1):170–6.
15. Chow E, Harth T, Hruba G, Finkelstein J, Wu J, Danjoux C. How accurate are physicians' clinical predictions of survival and the available prognostic tools in estimating survival times in terminally ill cancer patients? A systematic review. *Clin Oncol*. 2001;13(3):209–18.
16. Allen LA, Yager JE, Funk MJ, et al. Discordance between patient-predicted and model-predicted life expectancy among ambulatory patients with heart failure. *JAMA*. 2008;299(21):2533–42.
17. Selby D, Chakraborty A, Lilien T, Stacey E, Zhang L, Myers J. Clinician accuracy when estimating survival duration: the role of the patient's performance status and time-based prognostic categories. *J Pain Symptom Manage*. 2011;42(4):578–88.

18. Glare P, Virik K, Jones M, et al. A systematic review of physicians' survival predictions in terminally ill cancer patients. *BMJ*. 2003;327(7408):195–8.
19. Mendis R, Soo WK, Zannino D, Michael N, Spruyt O. Multidisciplinary prognostication using the Palliative Prognostic Score in an Australian cancer center. *Palliat Care*. 2015;9:7–14.
20. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. *JAMA*. 2012;307(2):182–92.
21. Chin MH, Drum ML, Jin L, Shook ME, Huang ES, Meltzer DO. Variation in treatment preferences and care goals among older patients with diabetes and their physicians. *Med Care*. 2008;46(3):275–86.
22. Robbins JM, Vaccarino V, Zhang H, Kasl SV. Socioeconomic status and type 2 diabetes in African American and non-Hispanic white women and men: evidence from the Third National Health and Nutrition Examination Survey. *Am J Public Health*. 2001;91(1):76–83.
23. Huang ES, Zhang Q, Gandra N, Chin MH, Meltzer DO. The effect of comorbid illness and functional status on the expected benefits of intensive glucose control in older patients with type 2 diabetes: a decision analysis. *Ann Intern Med*. 2008;149(1):11–9.
24. Clarke PM, Gray AM, Briggs A, et al. A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS No. 68). *Diabetologia*. 2004;47(10):1747–59.
25. 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):801–8.
26. Eastman RC, Javitt JC, Herman WH, et al. Model of complications of NIDDM. I. Model construction and assumptions. *Diabetes Care*. 1997;20(5):725–34.
27. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189–98.
28. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61(1):92–105.
29. Harrell FE. *Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer; 2001.
30. Graf E, Schmoor C, Sauerbrei W, Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. *Stat Med*. 1999;18(17–18):2529–45.
31. Gerds TA, Cai T, Schumacher M. The performance of risk prediction models. *Biom J*. 2008;50(4):457–79.
32. Huang ES, Sachs GA, Chin MH. Implications of new geriatric diabetes care guidelines for the assessment of quality of care in older patients. *Med Care*. 2006;44(4):373–7.
33. Samuelson W, Zeckhauser R. Status quo bias in decision making. *J Risk Uncertainty*. 1988;1:7–59.
34. Good CB. Polypharmacy in elderly patients with diabetes. *Diabetes Spectrum*. 2002;15(4):240–8.
35. Nelson JM, Dufraux K, Cook PF. The relationship between glycemic control and falls in older adults. *J Am Geriatr Soc*. 2007;55(12):2041–4.
36. Shorr RI, Ray WA, Daugherty JR, Griffin MR. Incidence and risk factors for serious hypoglycemia in older persons using insulin or sulfonylureas. *Arch Intern Med*. 1997;157(15):1681–6.
37. Hayes AJ, Leal J, Gray AM, Holman RR, Clarke PM. UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82. *Diabetologia*. 2013;56(9):1925–33.
38. Holmes HM. Rational prescribing for patients with a reduced life expectancy. *Clin Pharmacol Ther*. 2009;85(1):103–7.
39. Walter LC, Covinsky KE. Cancer screening in elderly patients: a framework for individualized decision making. *JAMA*. 2001;285(21):2750–6.
40. Christakis NA. *Death Foretold: Prophecy and Prognosis in Medical Care*. Chicago, IL: University of Chicago; 1999.
41. Goldstein MK, Hoffman BB, Coleman RW, et al. Implementing clinical practice guidelines while taking account of changing evidence: ATHENA DSS, an easily modifiable decision-support system for managing hypertension in primary care. *Proc AMIA Symp*. 2000:300–4.