University of Massachusetts Medical School

eScholarship@UMMS

University of Massachusetts Medical School Faculty Publications

Use of Risk Models to Predict Death in the Next Year Among Individual Ambulatory Patients With Heart Failure

Larry A. Allen
University of Colorado School of Medicine

Et al.

Let us know how access to this document benefits you.

Use of Risk Models to Predict Death in the Next Year Among Individual Ambulatory Patients With Heart Failure

Larry A. Allen, MD, MHS; Daniel D. Matlock, MD, MPH; Susan M. Shetterly, MS; Stanley Xu, PhD; Wayne C. Levy, MD; Laura B. Portalupi, MSW; Colleen K. McIlvennan, DNP, ANP; Jerry H. Gurwitz, MD; Eric S. Johnson, PhD, MPH; David H. Smith, RPh, MHA, PhD; David J. Magid, MD, MPH

IMPORTANCE The clinical practice guidelines for heart failure recommend the use of validated risk models to estimate prognosis. Understanding how well models identify individuals who will die in the next year informs decision making for advanced treatments and hospice.

OBJECTIVE To quantify how risk models calculated in routine practice estimate more than 50% 1-year mortality among ambulatory patients with heart failure who die in the subsequent year.

DESIGN, SETTING, AND PARTICIPANTS Ambulatory adults with heart failure from 3 integrated health systems were enrolled between 2005 and 2008. The probability of death was estimated using the Seattle Heart Failure Model (SHFM) and the Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC) risk calculator. Baseline covariates were collected from electronic health records. Missing covariates were imputed. Estimated mortality was compared with actual mortality at both population and individual levels.

MAIN OUTCOMES AND MEASURES One-year mortality.

RESULTS Among 10 930 patients with heart failure, the median age was 77 years, and 48.0% of these patients were female. In the year after study enrollment, 1661 patients died (15.9% by life-table analysis). At the population level, 1-year predicted mortality among the cohort was 9.7% for the SHFM (C statistic of 0.66) and 17.5% for the MAGGIC risk calculator (C statistic of 0.69). At the individual level, the SHFM predicted a more than 50% probability of dying in the next year for 8 of the 1661 patients who died (sensitivity for 1-year death was 0.5%) and for 5 patients who lived at least a year (positive predictive value, 61.5%). The MAGGIC risk calculator predicted a more than 50% probability of dying in the next year for 52 of the 1661 patients who died (sensitivity, 3.1%) and for 63 patients who lived at least a year (positive predictive value, 45.2%). Conversely, the SHFM estimated that 8496 patients (77.8%) had a less than 15% probability of dying at 1 year, yet this lower-risk end of the score range captured nearly two-thirds of deaths (n = 997); similarly, the MAGGIC risk calculator estimated a probability of dying of less than 25% for the majority of patients who died at 1 year (n = 914).

CONCLUSIONS AND RELEVANCE Although heart failure risk models perform reasonably well at the population level, they do not reliably predict which individual patients will die in the next year.

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Larry A. Allen, MD, MHS, Division of Cardiology, University of Colorado School of Medicine, Academic Office 1, Room 7019, 12631 E 17th Ave, Mail Stop B130, Aurora, CO 80045 (larry.allen@ucdenver.edu).

An accurate assessment of short-term mortality is central to the care of patients with heart failure. Implantable cardioverter-defibrillators are reserved for patients "who have a reasonable expectation of meaningful survival for more than 1 year."[1(pe183)] Left ventricular assist devices and transplantation are restricted to patients whose 1-year mortality without surgery exceeds the expected mortality with surgery (approximately 20% among patients with a durable left ventricular assist device).[2] Hospice eligibility relies on a physician's certification "that the patient's life expectancy is 6 months or less."[2(p32 176)]

Guidelines recommend using validated risk scores to estimate the risk of mortality for patients with heart failure,[1] and those "who are at high-risk for 1-year mortality using prognostic models should be referred for advanced therapy."[3(p159)] The Seattle Heart Failure Model (SHFM) is well-validated,[4-6] commonly used, and available online (https://depts.washington.edu/shfm/).[7] The Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC) risk calculator is validated, increasingly used, and available online (http://www.heartfailurerisk.org).[8] Yet, the application of population-risk data to individual treatment decisions has rarely been characterized.[9,10] We compared the SHFM and MAGGIC risk calculator estimates with actual deaths at 1 year among a cohort of ambulatory patients with heart failure, for the overall population and for individual patients.

Methods

All ambulatory patients 21 years of age or older with a diagnosis of heart failure[11,12] during the period from 2005 to 2008 were identified from Kaiser Permanente Colorado, Kaiser Permanente Northwest, and Fallon Health. Baseline covariates for the risk calculators (Table 1) were extracted from electronic health records. Deaths were identified from health plan databases, state death certificates, and Social Security Administration files.[13]

The SHFM and MAGGIC risk calculator scores were calculated using the online algorithms.[7,8] Mimicking the calculators, we imputed the mean values for missing data. New York Heart Association functional class, available in routine care but unavailable in electronic records, was set to functional class III in primary analysis and class IV in secondary analysis. To address concerns about model transportability, we updated the intercept and parameter estimates. The SHFM scores were converted to estimated survival at specific times using the following function: survival $(t) = \exp(-0.0405t)^{e^{(\text{SHFM score})}}$, where t is time. The MAGGIC risk calculator estimates for mortality were mapped based on probabilities for the integer scores 0 to 50, as described in the original derivation.[8] Following the published method,[8] we used multiple imputation for the left ventricular ejection fraction, body mass index, systolic blood pressure, serum creatinine level, and smoking status. The MAGGIC risk calculator uses the left ventricular ejection fraction in interactions with age and systolic blood pressure, so we performed a secondary analysis restricted to persons with left ventricular ejection fraction data. Actual survival was described using life-table analyses to handle truncated

Key Points

Question How well do heart failure risk models perform in identifying which patients will die in the next year?

Findings In this cohort study of 10 930 adults with heart failure, 1661 patients died in the next year. One risk model predicted a more than 50% probability of dying for only 8 of those patients, and another risk model predicted death for only 52 of those patients.

Meaning Heart failure risk models perform reasonably well at the population level, but they do not reliably predict which individual patients will die in the next year.

follow-up.[14] The institutional review boards at the 3 participating sites approved the study, and waiver of consent was obtained because of the nature of the study.

Results

Among 10 930 patients with heart failure, the median age was 77 years, and 48.0% of these patients were female (Table 1). In the year following registry inclusion, 1661 patients died (15.9% by life-table analysis).

Reflecting naturalistic data capture, 75% of the patients had at least 15 of the 20 SHFM covariates (range, 11-19 covariates). At the population level, the SHFM estimated that the 1-year mortality rate was 9.7%, with a median life expectancy of 7.0 years (interquartile range, 5.3-9.0 years; range, 3 months [shortest predicted remaining survival of anyone in the cohort] to 22 years [longest]). Discrimination was similar to other cohorts: the C statistic was 0.69 at 6 months, 0.66 at 1 year, and 0.65 at 3 years. Model calibration showed underestimations: a 6-month predicted mortality rate of 5.0% vs an observed mortality rate of 10.5%; a 1-year predicted mortality rate of 9.7% vs an observed mortality rate of 15.9%; and a 3-year predicted mortality rate of 24.6% vs an observed mortality rate of 33.6%. The Hosmer-Lemeshow test suggested an adequate model fit ($\chi^2 = 8.7$, $P = .36$).

At the individual level, the SHFM performed poorly (Figure). The SHFM predicted a more than 50% mortality rate in the next year for 8 of the 1661 patients who actually died (sensitivity for 1-year death was 0.5%) and for 5 patients who lived at least a year (positive predictive value, 61.5%) (Table 2). In secondary analysis forcing the New York Heart Association to the worst functional class, the SHFM predicted a more than 50% 1-year mortality rate among 49 patients who died (sensitivity, 3.0%) and among 47 patients who lived at least a year (positive predictive value, 51.0%). Updating the SHFM intercept and parameter estimates to recalibrate and refit the model (which could not be done in clinical practice), we found that the C statistic improved (0.75) but still captured less than 8% of patients who died in the next year (of 191 patients predicted to die, 123 actually died). Conversely, the SHFM estimated that 8496 patients (77.7%) had a more than 85% probability of surviving 1 year, yet this "lower-risk" end of the score range captured nearly two-thirds of deaths (n = 997).

Table 1. Baseline Characteristics Used in the SHFM and the MAGGIC Score Calculation

Variable	Patients, No. (%) (N = 10 930)	Covariate SHFM	MAGGIC
Age, mean [SD], y	75.1 [11.8]	x	x
Male sex	5681 (52.0)	x	x
NYHA class[a]		x	x
Weight,[b] mean [SD], kg	84.8 [24.2]	x	
BMI			x
<15	10 (0.1)		
15-19	336 (3.1)		
20-24	1776 (16.2)		
25-29	2629 (24.1)		
≥30	3480 (31.8)		
Missing	2699 (24.7)		
LVEF category		x	x
Preserved (≥50%)	4155 (38.0)		
Borderline (41%-49%)	1330 (12.2)		
Reduced (≤40%)	3019 (27.6)		
Missing	2426 (22.2)		
SBP, mean [SD], mm Hg	129.2 [22.9]	x	x
<110	1524 (13.9)		
110-119	1389 (12.7)		
120-129	1739 (15.9)		
130-139	1511 (13.8)		
140-149	1203 (11.0)		
≥150	1503 (13.8)		
Missing	2061 (18.9)		
Ischemic etiology	2239 (20.5)	x	
Prescription drug use			
ACE inhibitor	4499 (41.2)	x	
ARB	1303 (11.9)	x	
ACE inhibitor/ARB	5705 (52.2)		x
β-Blocker	6374 (58.3)	x	x
Statin	4734 (43.3)	x	
Allopurinol	637 (5.8)	x	
Aldosterone blocker	1106 (10.1)	x	
Furosemide equivalents,[c] mg (24-h dose)		x	
No loop diuretic Rx	5471 (50.1)		
<20	16 (0.1)		
20	2287 (20.9)		
27[d]	127 (1.2)		
40	2474 (22.6)		
53[d]	156 (1.4)		
80	286 (2.6)		
100	84 (0.8)		
200	29 (0.3)		

(continued)

Table 1. Baseline Characteristics Used in the SHFM and the MAGGIC Score Calculation (continued)

Variable	Patients, No. (%) (N = 10 930)	Covariate SHFM	MAGGIC
Laboratory data			
Hemoglobin,[e] mean [SD], g/dL	13.0 [1.9]	x	
Lymphocyte,[d,f] mean [SD], %	22.3 [9.9]	x	
Uric acid,[d,g] mean [SD], mg/dL	7.54 [2.3]	x	
Total cholesterol,[h] mean [SD], mg/dL	155.7 [42.3]	x	
Sodium,[i] mean [SD], mEq/L	139.1 [3.8]	x	
Serum creatinine, mg/dL			x
<1.02	4032 (36.9)		
1.02-1.23	2127 (19.5)		
1.24-1.46	1386 (12.7)		
1.47-1.69	889 (8.1)		
1.70-1.91	632 (5.8)		
1.92-2.36	394 (3.6)		
2.37-2.82	206 (1.9)		
≥2.83	388 (3.5)		
Missing	876 (8.0)		
Device			
ICD	632 (5.8)	x	
CRT	583 (5.3)	x	
Comorbidity			
Diabetes	1812 (16.6)		x
Lung disease[j]	3957 (36.2)		x
HF diagnosis ≤18 mo	6394 (58.5)		x
Current smoker[k]	390 (3.6)		x

Abbreviations: ACE, angiotensin-converting enzyme; ARB, angiotensin II receptor blocker; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CRT, cardiac resynchronization therapy; HF, heart failure; ICD, implantable cardioverter-defibrillator; MAGGIC, Meta-Analysis Global Group in Chronic Heart Failure; NYHA, New York Heart Association; LVEF, left ventricular ejection fraction; Rx, prescription; SBP, systolic blood pressure; SHFM, Seattle Heart Failure Model.

SI conversion factors: To convert hemoglobin to grams per liter, multiply by 10.0; uric acid to micromoles per liter, multiply by 59.485; total cholesterol to millimoles per liter, multiply by 0.0259; sodium to millimoles per liter, multiply by 1.0; serum creatinine to micromoles per liter, multiply by 88.4.

[a] Missing data for 10 930 patients (100%).

[b] Weight in kilograms is not directly in the SHFM model but is used to calculate furosemide dose per kilograms per day. Weight was missing for 16.1% of patients receiving furosemide and for 2148 patients (19.7%) in our study.

[c] Bumetanide (dose equivalent conversion 3 mg = 80 mg of furosemide); toresmide (dose equivalent conversion 1 mg = 2 mg of furosemide).

[d] Data only collected from Kaiser Permanente Colorado; the other 2 sites used default for 100% of medical records.

[e] Missing data for 2043 patients (18.7%).

[f] Missing data for 8453 patients (77.3%).

[g] Missing data for 10 583 patients (96.8%).

[h] Missing data for 3383 patients (31.0%).

[i] Missing data for 1384 patients (12.7%).

[j] The MAGGIC risk calculator uses the presence or absence of chronic obstructive pulmonary disease; an only slightly broader variable for pulmonary disease is available in these data.

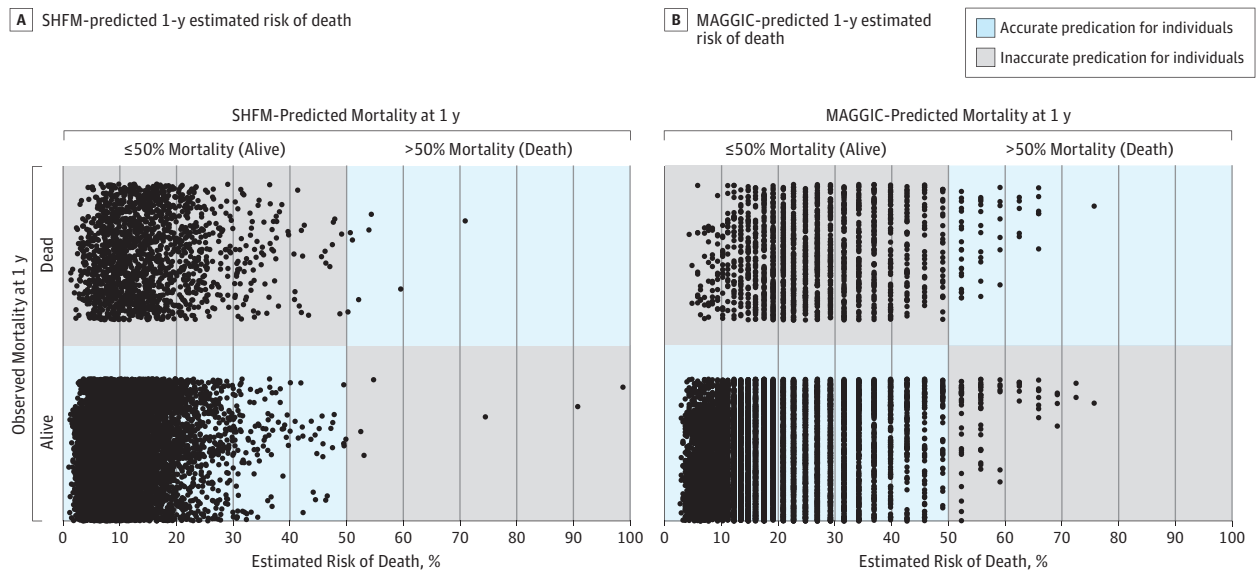[k] Missing data for 6916 patients (63.3%).

For the more stringent hospice referral threshold, the SHFM identified only 2 patients with a more than 50% probability of dying within 6 months, neither of whom died (5 patients had a New York Heart Association functional class set to IV). For a lower threshold of less than 20% 1-year mortality (which can guide a durable left ventricular assist device), the SHFM showed a sensitivity of 20.7%, a specificity of 93.1%, a positive predictive value of 39.6%, and a negative predictive value of 84.4%.

Analysis was repeated using the MAGGIC risk calculator. At the population level, the median score was 26 with a pre-dicted 1-year mortality rate of 17.5%. The C statistic at 1 year was 0.69. The Hosmer-Lemeshow statistic had a χ^2 of 38.6 ($P < .001$), indicating some lack of fit in this population. In secondary analyses among the 8504 patients with available left ventricular ejection fraction data, overall risk of death was lower (1150 died [13.5%]), the C statistic was 0.70, and the Hosmer-Lemeshow χ^2 was 16.9 ($P = .02$).

At the individual level, the MAGGIC risk calculator also performed poorly (Figure). It predicted a more than 50% mortality rate in the next year for 52 of the 1661 patients who actually died (sensitivity, 3.1%) and for 63 patients who lived

Figure. One-Year Estimated Risk of Death and Actual Death

A SHFM-predicted 1-y estimated risk of death

B MAGGIC-predicted 1-y estimated risk of death

☐ Accurate predication for individuals
☐ Inaccurate predication for individuals

Shown are the 1-year predictions of death for individuals and associated actual survival using the Seattle Heart Failure Model (SHFM) (A) and the Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC) risk calculator (B). Each dot represents an individual.

Table 2. Patients Categorized by Model-Predicted and Actual Mortality at 1 Year

| Model | 1-y Vital Status | | No./Total No. (%) |
	Died	Lived	
SHFM			
Predicted survival ≤50% (anticipated dead), No.	8	5	8/13 (61.5)[a]
Predicted survival >50% (anticipated alive), No.	1653	7616	7616/9269 (82.2)[b]
No./total No. (%)	8/1661 (0.5)[c]	7616/7621 (99.9)[d]	9282[e]
MAGGIC risk calculator			
Predicted survival ≤50% (anticipated dead), No.	52	63	52/115 (45.2)[a]
Predicted survival >50% (anticipated alive), No.	1609	7558	7558/9167 (82.4)[b]
No./total No. (%)	52/1661 (3.1)[c]	7558/7621 (99.2)[d]	9282[e]

Abbreviations: MAGGIC, Meta-Analysis Global Group in Chronic Heart Failure; SHFM, Seattle Heart Failure Model.

[a] Positive predictive value.

[b] Negative predictive value.

[c] Sensitivity.

[d] Specificity.

[e] A total of 1648 patients were censored owing to a lack of 1-year follow-up (25 patients with a MAGGIC-predicted survival of ≤50%).

at least a year (positive predictive value, 45.2%) (Table 2). Conversely, the MAGGIC risk calculator estimated that 8161 patients (74.7%) had a more than 75% probability of surviving 1 year, yet this "lower-risk" end of the score range captured the majority of deaths (n = 914). For the lower threshold of less than 20% 1-year mortality, the MAGGIC risk calculator showed a sensitivity of 69.7%, a specificity of 61.2%, a positive predictive value of 28.1%, and a negative predictive value of 90.3%.

Discussion

In this cohort of more than 10 000 ambulatory patients with heart failure, well-accepted risk models assigned a more than 50% chance of death in the next year to a very small fraction of those who actually died. Conversely, the majority of patients who died had a more than 75% model-estimated probability of surviving beyond 1 year. This highlights the limita-tions of applying population-based prognostic model results to short-term medical decisions about individual patients[15,16] and may partially justify the limited uptake of risk models in routine clinical practice.[17]

Most research has gone into refining risk models despite diminishing improvements in performance. Meanwhile, there is a dearth of information to help patients and health care professionals manage the uncertainty inherent in survival estimates.[18] Objective risk information should be integrated with recognition of uncertainty.[15,19,20] For example, prediction intervals may be more informative than median values.[21] Rather than say that "your expected survival is about 1 year," a clinician could say "if I had 100 people like you, 25 would die within 6 months and 25 would be alive after 2 years."

Limitations

Limitations should be considered. First, missing data requiring imputation may significantly relate to the poor risk pre-

diction seen in these models; however, imputation replicates most real-world risk assessments. Second, the model calculation was performed some years ago, but the rates of therapy and mortality for heart failure remained relatively unchanged from 2005 to 2014.[22,23] Finally, the risk models were developed in cohorts that either reflected secular differences in time (the SHFM) or geographical differences in care (the MAGGIC risk calculator). Moreover, the cohort used here—like any insured population in an integrated health plan—may fail to track well with the general population. Variability in the performance of the SFHM and the MAGGIC risk calculator is to be expected because the derivation cohorts came from highly orchestrated clinical trial populations instead of unselected general care patients.

Conclusions

Risk models cannot prospectively identify the vast majority of individual ambulatory patients who will die in the near future. For undulating diseases such as heart failure, improvements in risk modeling are only a partial solution. Meanwhile, we must embrace the certainty of uncertainty and better help patients plan for the worst while hoping for the best.

REFERENCES

1. Yancy CW, Jessup M, Bozkurt B, et al; American College of Cardiology Foundation; American Heart Association Task Force on Practice Guidelines. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* 2013;62(16):e147-e239.

2. Centers for Medicare & Medicaid Services (CMS), HHS. Medicare and Medicaid programs: hospice conditions of participation: final rule. *Fed Regist.* 2008;73(109):32087-32220.

3. Feldman D, Pamboukian SV, Teuteberg JJ, et al; International Society for Heart and Lung Transplantation. The 2013 International Society for Heart and Lung Transplantation Guidelines for mechanical circulatory support: executive summary. *J Heart Lung Transplant.* 2013;32(2):157-187.

4. Alba AC, Agoritsas T, Jankowski M, et al. Risk prediction models for mortality in ambulatory patients with heart failure: a systematic review. *Circ Heart Fail.* 2013;6(5):881-889.

5. May HT, Horne BD, Levy WC, et al. Validation of the Seattle Heart Failure Model in a community-based heart failure population and enhancement by adding B-type natriuretic peptide. *Am J Cardiol.* 2007;100(4):697-700.

6. Kalogeropoulos AP, Georgiopoulou VV, Giamouzis G, et al. Utility of the Seattle Heart Failure Model in patients with advanced heart failure. *J Am Coll Cardiol.* 2009;53(4):334-342.

7. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation.* 2006;113(11):1424-1433.

8. Pocock SJ, Ariti CA, McMurray JJ, et al; Meta-Analysis Global Group in Chronic Heart Failure. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J.* 2013;34(19):1404-1413.

9. O'Callaghan A, Laking G, Frey R, Robinson J, Gott M. Can we predict which hospitalised patients are in their last year of life? a prospective cross-sectional study of the Gold Standards Framework Prognostic Indicator Guidance as a screening tool in the acute hospital setting. *Palliat Med.* 2014;28(8):1046-1052.

10. Haga K, Murray S, Reid J, et al. Identifying community based chronic heart failure patients in the last year of life: a comparison of the Gold Standards Framework Prognostic Indicator Guide and the Seattle Heart Failure Model. *Heart.* 2012;98(7):579-583.

11. Go AS, Magid DJ, Wells B, et al. The Cardiovascular Research Network: a new paradigm for cardiovascular quality and outcomes research. *Circ Cardiovasc Qual Outcomes.* 2008;1(2):138-147.

12. Go AS, Yang J, Gurwitz JH, Hsu J, Lane K, Platt R. Comparative effectiveness of different beta-adrenergic antagonists on mortality among adults with heart failure in clinical practice. *Arch Intern Med.* 2008;168(22):2415-2421.

13. Go AS, Yang J, Ackerson LM, et al. Hemoglobin level, chronic kidney disease, and the risks of death and hospitalization in adults with chronic heart failure: the Anemia in Chronic Heart Failure: Outcomes and Resource Utilization (ANCHOR) Study. *Circulation.* 2006;113(23):2713-2723.

14. Collett D. *Modelling Survival Data in Medical Research.* New York, NY: Chapman & Hall; 1994.

15. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. *Med Decis Making.* 2007;27(5):681-695.

16. Kent DM, Shah ND. Risk models and patient-centered evidence: should physicians expect one right answer? *JAMA.* 2012;307(15):1585-1586.

17. Steyerberg EW, Moons KGM, van der Windt DA, et al; PROGRESS Group. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med*. 2013;10(2):e1001381.

18. Smith AK, White DB, Arnold RM. Uncertainty—the other side of prognosis. *N Engl J Med*. 2013;368(26):2448-2450.

19. Henderson R, Jones M, Stare J. Accuracy of point predictions in survival analysis. *Stat Med*. 2001;20(20):3083-3096.

20. Henderson R, Keiding N. Individual survival time prediction using statistical models. *J Med Ethics*. 2005;31(12):703-706.

21. Olive DJ. Prediction intervals for regression models. *Comput Stat Data Anal*. 2007;51(6):3115-3122. doi:10.1016/j.csda.2006.02.006

22. Sacks CA, Jarcho JA, Curfman GD. Paradigm shifts in heart-failure therapy—a timeline. *N Engl J Med*. 2014;371(11):989-991.

23. Mozaffarian D, Benjamin EJ, Go AS, et al; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. *Circulation*. 2015; 131(4):e29-e322.