The Art versus Science of Predicting Prognosis: Can a Prognostic Index Predict Short-Term Mortality Better than Experienced Nurses Do?

David J. Casarett, M.D., M.A.,[1] Sue Farrington, M.B.A.,[2] Teresa Craig, C.P.A.,[2] Julie Slattery, M.S., R.N.,[3]
Joan Harrold, D.O.,[4] Betty Oldanie, R.N., B.S.N., M.A.,[5] Jason Roy, Ph.D.,[6]
Richard Biehl, Ph.D.,[7] and Joan Teno, M.D., M.S.[8]

Abstract

Objective: To determine whether a prognostic index could predict one-week mortality more accurately than hospice nurses can.

Method: An electronic health record-based retrospective cohort study of 21,074 hospice patients was conducted in three hospice programs in the Southeast, Northeast, and Midwest United States. Model development used logistic regression with bootstrapped confidence intervals and multiple imputation to account for missing data. The main outcome measure was mortality within 7 days of hospice enrollment.

Results: A total of 21,074 patients were admitted to hospice between October 1, 2008 and May 31, 2011, and 5562 (26.4%) died within 7 days. An optimal predictive model included the Palliative Performance Scale (PPS) score, admission from a hospital, and gender. The model had a c-statistic of 0.86 in the training sample and 0.84 in the validation sample, which was greater than that of nurses' predictions (0.72). The index's performance was best for patients with pulmonary disease (0.89) and worst for patients with cancer and dementia (both 0.80). The index's predictions of mortality rates in each index category were within 5.0% of actual rates, whereas nurses underestimated mortality by up to 18.9%. Using the optimal index threshold (<3), the index's predictions had a better c-statistic (0.78 versus 0.72) and higher sensitivity (74.4% versus 47.8%) than did nurses' predictions but a lower specificity (80.6% versus 95.1%).

Conclusions: Although nurses can often identify patients who will die within 7 days, a simple model based on available clinical information offers improved accuracy and could help to identify those patients who are at high risk for short-term mortality.

Introduction

THE U.S. MEDICARE HOSPICE BENEFIT was established in 1982 to ensure that patients in the last 6 months of life have access to high-quality palliative care.[1] Since 1982, the hospice industry has grown rapidly. In 2009, approximately 1,560,000 people used hospice in the United States, compared with 1,200,000 in 2005.[2]

Although patients are seeking hospice care in increasing numbers, they are still enrolling in hospice very late in the course of illness. For instance, half of hospice patients are still referred in the last 3 weeks of life.[2] Moreover, one-third of patients are referred in the last week and 10% are referred in the last 24 hours. Patients with shorter lengths of stay,[3] and particularly those whose families say they were referred too late,[4] may have more unmet needs for care.

It is possible that improved advance care planning and other interventions could encourage patients to consider hospice earlier. However, it also will be be essential to explore ways in which hospices can more effectively meet the needs of

[1]Departments of Medicine and [6]Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania.
[2]Suncoast Solutions Inc., Clearwater, Florida.
[3]Agrace HospiceCare, Madison, Wisconsin.
[4]Hospice of Lancaster County, Lancaster, Pennsylvania.
[5]Suncoast Hospice, Clearwater, Florida.
[7]University of South Florida, Tampa, Florida.
[8]Warren Alpert School of Medicine, Brown University, Providence, Rhode Island.

patients who are referred to hospice very late. That is, it will be important to ensure that these patients receive aggressive "front-loaded" services. Moreover, some patients who are at high risk for dying very soon and who have severe symptoms could receive care in an inpatient unit that is optimally designed to provide high-intensity palliative care.

It is not known how accurately hospice nurses—the de facto "gold standard"—are able to identify patients who are likely to die within 7 days of enrollment. Nor is it known whether a prognostic model can improve the accuracy of their predictions. Therefore, the goals of this study were to determine how well hospice providers are able to identify patients at high risk for dying within one week, and to determine whether a simple prognostic index could improve these predictions.

Method

Setting and sample

The three participating hospices (Suncoast Hospice, Clearwater, FL; Agrace HospiceCare, Madison WI; Hospice of Lancaster County, Lancaster, PA) are pilot members of the CHOICE network (Coalition of Hospices Organized to Investigate Comparative Effectiveness). CHOICE hospices all use Suncoast Solutions Inc.'s Electronic Health Record (EHR) Software, and all agree to permit use of their data for research projects that are defined by the coalition's members. Participating hospices range in size (daily census) from 500 to 2700 patients/day. Data elements for this study were defined a priori and then extracted from participating hospices' EHRs. Extracted data were then stripped of identifiers to create a Health Insurance Portability and Accountability Act (HIPAA)-compliant limited dataset, and transferred as an encrypted file to the University of Pennsylvania for analysis.

Data analysis

We used logistic regression models to identify variables that were associated with an increased likelihood of dying within 7 days. In selecting candidate variables, we focused on those that would be readily available on admission, to develop a model that hospice providers and referring physicians could apply at the time of a patient's enrollment in hospice. We used a limited set of demographic variables, including age, gender, race (white versus nonwhite), site of care (e.g., hospital, nursing home) as well as marital status, which has been described as a predictor of hospice survival.[5]

As a measure of functional status, we included the Palliative Performance Scale (PPS), which has been widely studied as a predictor of survival in hospice and palliative care populations.[5–8] The PPS is scored from 0 to 100 in 10-point increments to create an 11-point scale in which higher numbers correspond to better functional status. We also considered clinical variables that have not been studied as predictors of prognosis in this population, but which are plausibly associated with poor survival. These included the presence of a wound (venous stasis, pressure ulcer, or malignant), the use of medical devices (e.g., intravenous or urinary catheters), and the location of care at the time of hospice enrollment (hospital versus other).

All continuous variables were coded as is, to preserve predictive power,[9] with two exceptions. First, because of small cell sizes at the higher (better) end of the PPS scale, we recoded the PPS into six categories (0–10, 20, 30, 40, 50, and 60–100) based on previous studies.[7] Second, the presence of many uncommon diagnoses made it more efficient to group diagnoses into the top six categories (five diagnoses plus "other").

We had decided a priori to exclude variables for which more than 10% of values were missing or out of range. However, rates for all variables under consideration were well below this limit (range: 0.02%–4.1%, interquartile range: 0.4%–3.2%). Therefore, we restricted prognostic models by using a casewise deletion of observations with missing variables.[10–13]

We developed a model and tested the resulting survival index in sequential steps.[14,15] First, we created two independent samples of approximately equal size by selecting the two smaller hospices for model training ($n = 8496$). We then used the remaining hospice for model validation ($n = 12,578$). This is consistent with current recommendations of prognostic model development that suggest development and validation in distinct and independent populations.[14,16] In the model development sample, we examined each of the variables in Table 1 as potential predictors of death within 7 days.[17,18]

Because the number of variables under consideration was small, we used backwards selection to maximize the predictive value of the resulting model.[15,19] For each version, we used bootstrapped confidence intervals to provide internal validation of the c-statistic.[20] We reasoned that an index with as few variables as possible would be most clinically useful, so we restricted candidate models to those in which all variables were significant. From these, we selected the final model based on its c-statistic and optimal Bayes Information Criterion (BIC), a modification of the Akaike Information Criterion.[15,21]

Next, we repeated the steps above using multiple imputation to assess the impact of missing data on the model's fit. We then compared the imputed model derived from the entire training sample with the model derived from the subset of patients with complete data. To assess the model's optimism, or potential for inflated estimates of accuracy, we tested the final model in the validation sample, again using bootstrapped confidence intervals. We then compared the original and validation c-statistics. In general, a large reduction in the c-statistic in the validation sample suggests that its predictions may not be generalizable.

For each patient in the validation sample with complete data, we calculated a prognostic score corresponding to the absolute value of the sum of his/her coefficients (the prognostic index). We assessed the index's predictive validity in three ways. First, we calculated the predicted and actual one-week mortality rates for each index level. Second, we selected the index score with the best c-statistic and calculated its sensitivity, specificity, and positive and negative predictive values. Third, to examine the index's calibration, we used the binomial test of proportions to examine differences in the number of observed versus expected deaths among the index groups.

Finally, we examined the accuracy of nurses' predictions of a patient's survival. To do this, we used responses to a question embedded in the EHR admission form that asks the admitting nurse whether death is "imminent." Although this term is not defined in the EHR, we reasoned that it would offer a good assessment of a patient's short-term prognosis. We calculated a c-statistic and test characteristics of nurses' predictions, and compared these predictions with actual results using the binomial test. We also tested a hybrid model

TABLE 1. CHARACTERISTICS OF PATIENTS ALIVE VERSUS THOSE WHO DIED WITHIN 7 DAYS OF HOSPICE ADMISSION

	N=21,074 (entire sample)	Died within 7 days (n=5562)	Alive after 7 days (n=15,512)	β coefficient (bivariate), 95% confidence interval, P value
Age: mean (interquartile range)	79.3 (72–89)	80.1 (74–89)	79.1 (72–89)	0.006 (0.003, 0.008) $p < 0.001$
Female	11,772 (55.9%)	2958 (53.2%)	8814 (56.8%)	−0.15; (-0.21, −0.86); $p < 0.001$
Race: n (%)				
White	19,045 (90.4%)	5004 (90.0%)	14,041 (90.5%)	0.013; (-0.12, −0.15); $p = 0.838$
Nonwhite	1192 (5.7%)	310 (5.6%)	882 (5.7%)	
Missing	837 (4.0%)	248 (4.5%)	589 (3.8)	
Married: n (%)	6423 (30.5%)	1677 (30.2%)	4746 (30.6%)	−0.02; (-0.09, −0.05); $p = 0.537$
Location at admission: home or long-term care[a]	15,221 (72.2%)	2693 (48.2%)	12,528 (80.8%)	−1.50; (-1.56, −1.43); $p < 0.001$
Palliative Performance Scale (first 24 hours)				
0–10	2225 (10.6%)	1779 (32.0%)	446 (2.9%)	—
20	2196 (10.4%)	1283 (23.1%)	913 (5.9%)	−1.04; (-1.18, −0.91); $p < 0.001$
30	4101 (19.4%)	1171 (21.0%)	2930 (18.9%)	−2.30; (-2.42, −2.18); $p < 0.001$
40	6473 (30.7%)	747 (13.4%)	5726 (36.9%)	−3.42; (-3.55, −3.30); $p < 0.001$
50	3867 (18.4%)	234 (42.%)	3633 (23.4%)	−4.12; (-4.29, −3.96); $p < 0.001$
60–100	1005 (4.8%)	51 (0.9%)	954 (6.2%)	−4.31; (-4.61, −4.01); $p < 0.001$
PPS score done >24 hours postadmission	345 (1.6)	8 (0.1%)	337 (2.2%)	
Missing	862 (4.1%)	289 (52%)	573 (3.7%)	
Foley catheter or intermittent catheterization	5886 (27.9%)	2322 (41.8%)	3564 (23.0%)	0.88; (0.81, 0.94); $p < 0.001$
Oxygen	7491 (35.6%)	2985 (53.7%)	4506 (29.0%)	1.04; (0.98, 1.10); $p < 0.001$
Gastrostomy or jejunostomy tube	579 (2.8%)	119 (2.1%)	460 (3.0%)	−0.33; (-0.54, −0.13); $p = 0.001$
Wound (malignant or pressure ulcer)	2657 (12.6%)	442 (8.0%)	2215 (14.3%)	−0.66; (-0.76, −0.55); $p < 0.001$
Primary diagnosis				
Cancer	7391 (35.1%)	1499 (27.0%)	5892 (38.0%)	−0.92; (-1.02, −0.02); $p < 0.001$
Debility	3220 (15.3%)	709 (12.8%)	2511 (16.2%)	−0.81; (-0.93, −0.70); $p < 0.001$
Cardiovascular disease	2968 (14.1%)	796 (14.3%)	2172 (14.0%)	−0.55; (-0.67, −0.44); $p < 0.001$
Dementia	2521 (12.0%)	635 (11.4%)	1886 (12.2%)	−0.64; (-0.76, −0.52); $p < 0.001$
Pulmonary disease	1414 (6.7%	434 (7.8%)	980 (6.3%)	−0.36; (-0.50, −0.22); $p < 0.001$
Stroke	1016 (4.8%)	499 (9.0%)	517 (3.3%)	−0.42; (-0.26, −0.56); $p < 0.001$
Other	2544 (12.1%)	990 (17.8%)	1554 (10%)	—

[a]Includes: patient's own home, friend/family member's home, assisted living, group home, prison, homeless shelter.

that incorporated nurses' predictions and the prognostic index, comparing its accuracy with either alone.

We estimated that a training sample of at least 8000 patients with a one-week mortality rate of 25% would provide adequate power to detect a 5% increase (e.g., between 25% and 30%) in the risk of one-week mortality attributable to a single variable with a prevalence of 10% ($\alpha = 0.05$). Although no standards exist for the estimation of necessary sample sizes for logistic regression models, a rough rule of thumb is to allow at least 10 events for each variable under consideration.[22,23] Therefore, this sample would provide adequate power to detect even small effects of a candidate predictor variable on the risk of one-week mortality.

The University of Pennsylvania's Institutional Review Board approved the use of secondary data for this study. Stata statistical software (Stata MP2 11.0 for Mac, StataCorp., College Station, TX) was used for all statistical analysis.

Results

A total of 21,074 patients were admitted between October 1, 2008 and May 31, 2011. Hospices contributed 4475 patients (Lancaster), 4021 patients (Madison), and 12,578 patients (Clearwater). Of all patients admitted during the study period, 5562 (26.4%) died within 7 days. An additional 122 patients who were discharged alive from hospice in the first 7 days were assumed to have lived at least 7 days. Seven-day mortality rates were similar across the three hospices (Lancaster 27.3%, Madison 25.4%, Clearwater 26.4%).

The characteristics of the sample are described in Table 1. Compared with national descriptions of the U.S. hospice population,[2] this sample had a slightly longer length of stay (median 25 days versus 21 days), and patients in this sample were less likely to die within 7 days (26.4% versus 33%). This sample resembled the national hospice population with respect to gender (56% versus 54% female), the proportion with cancer (37% versus 41%), and age (84% versus 83% >65). The sample had a higher prevalence of white patients compared with the national hospice population (96% versus 80%), but was typical of the populations of these three regions.

Model development and validation

Of 21,074 patients, 8496 (40.3%) were from the hospices used for model training, and 12,578 (59.7%) were from the remaining hospice used for model validation. One-week mortality rates were identical in the development (2242/8496; 26.4%) and validation samples (9258/12,578; 26.4%).

Most patients (19,867; 94.3%) had a PPS score assessed within 24 hours of admission, but 345 (1.6%) had a score assessed after 24 hours, and 862 (4.1%) never had a score documented. Patients without a PPS score were more likely to die within 7 days (289/862; 33.5%) compared with patients with an early PPS score (5265/19,867; 26.5%). Conversely, patients with a late PPS score were less likely to die within 7 days (8/345; 2.3%). Because these late PPS patients necessarily had a longer survival, we excluded them from analysis.

Backwards selection using all variables in Table 1 generated an optimal model that was chosen based on the inclusion of only significant variables and an optimum c-statistic (0.83) and BIC (6427.47). Variables included in the model are displayed in Table 2. This model was based on 7425 patients in the development sample with complete data, out of a possible 8496 (87%). We followed the same process using multiple imputation, but because the resulting model was virtually identical, the nonimputed (casewise deletion) model was used. The final model was run again with bootstrapping on the validation sample with a c-statistic of 0.84, compared with a c-statistic in the development sample of 0.83, indicating no optimism of the original model.

Next, we calculated a survival index for each patient in the training sample, using the sum of the absolute value of each patient's coefficients. For instance, a female patient admitted from home with a PPS score of 20 had a score of 1.93 (1.02 [PPS score] + 0.48 [home] + 0.43 [gender]), which would be rounded to an index score of 2. The resulting scores ranged from 0 to 5 and higher numbers corresponded to a better prognosis (lower probability of 7-day mortality). The index's overall c-statistic was 0.86, and the best c-statistic was for an index < 3 (0.75). This was considerably better than the most informative variable in the model (PPS score), which had a c-statistic of only 0.74 overall, and 0.67 for a score of < 40.

We then calculated an index for each patient in the validation sample. The index had a c-statistic in the validation sample of 0.84 for those with complete data ($n = 12,442$). As in the development sample, the optimal c-statistic corresponded to a score of < 3 (0.78). This threshold had a sensitivity of 74.4%, a specificity of 80.6%, and positive and negative predictive values of 57.9% and 89.8%, respectively.

Because diagnosis was not included in the final model, we examined the index's performance across the most common

Table 2. Final Prognostic Model Predicting 7-day Mortality among Hospice Patients

	β coefficient	95% confidence interval	P value
Female	−0.43	−0.78, −0.07	$p = 0.020$
Location at admission: Home or long-term care versus all others	−0.48	−0.75, −0.22	$p < 0.001$
Palliative Performance Scale score			
0–10	—		
20	−1.02	−1.61, −0.43	$p < 0.001$
30	−2.08	−2.62, −1.54	$p < 0.001$
40	−2.86	−3.49, −2.23	$p < 0.001$
50	−3.85	−4.56, −3.15	$p < 0.001$
60–100	−4.22	−5.31, −3.13	$p < 0.001$

diagnoses (Table 1). The c-statistic was highest for pulmonary disease (0.89). Other diagnoses had lower c-statistics, including cardiovascular disease (0.85), debility (0.85), stroke (0.82), cancer (0.80), and dementia (0.80).

Nurses' predictions

Of 12,578 patients in the validation sample, nurses predicted an imminent death for 2074 (16.5%). The c-statistic for predictions of death at various time periods between one and 10 days ranged from 0.65 (one day), 0.64 (3 days), 0.72 (7 days), 0.63 (10 days), and 0.61 (14 days). Patients for whom death was believed to be imminent were much more likely to die within 7 days (1617/2074 versus 1703/10,504; 78.0% versus 16.2%; odds ratio: 18.28; 95% confidence interval: 16.3–20.5; $p < 0.001$). A prediction of imminent death had a c-statistic of 0.72, a sensitivity of 48.7%, a specificity of 95.1%, and positive and negative predictive values of 78.0% and 83.8%, respectively. Thus compared with the model's predictions at a threshold of < 3, nurses' predictions offered a lower sensitivity (47.8% versus 74.4%) but a higher specificity (95.1% versus 80.6%) and overall worse c-statistic (0.72 versus 0.78).

Compared with nurses' predictions, the index offered better calibration across all index levels (Fig. 1). For a score of 5, the 7-day mortality (45/1242; 3.6%) was greater than nurses' predictions (1.3%) but less than the index's prediction (4.6%). For lower (worse) index scores, nurses' predictions increasingly underestimated the probability of 7-day mortality. For instance, for patients with an index of 4 (261/3351; 7.8%), the predicted mortality was 10.5% but nurses' was 2.3%. For an index score of 3 (533/3634; 14.7%), the predicted mortality was 19.7% but nurses' was 4.7%. For an index score of 2 (684/1867; 36.6%), the predicted mortality was 36.8% and nurses' was 17.9%. For patients with an index score of 1 (843/1241; 67.9%) the predicted mortality (63.1%) was higher than the nurses' predicted 49.0% mortality. Finally, for patients with an index score of 0 (912/1107; 82.4%), the predicted mortality (81.4%) was also higher than the nurses' predictions (75.4%).

Because the model's and nurses' predictions offered greater sensitivity and specificity, respectively, we developed a hybrid index that incorporated both. We created five versions by adding between 1 and 10 points if death was not coded as being imminent. In the validation sample, none of these hybrid models offered a c-statistic that was greater than that of the original index. The optimal index was obtained when an additional 5 points were assigned when death was not noted to be imminent. This index had a c-statistic of 0.84, which was the same as that of the original index.

Discussion

Despite rapid growth in the hospice industry, patients still enroll very late in the course of illness. Therefore, hospices and patients' health care providers will need to put systems in place that allow them to meet the needs of patients who enroll in the last days of life. This study provides two key findings that will guide these efforts.

First, this study found that hospice nurses are reasonably accurate in predicting 7-day mortality. This is unexpected because previous studies have found that health care providers are poor at predicting prognosis,[24,25] and even palliative care providers tend to be optimistic.[26] However, these results are more in line with reports of providers working in
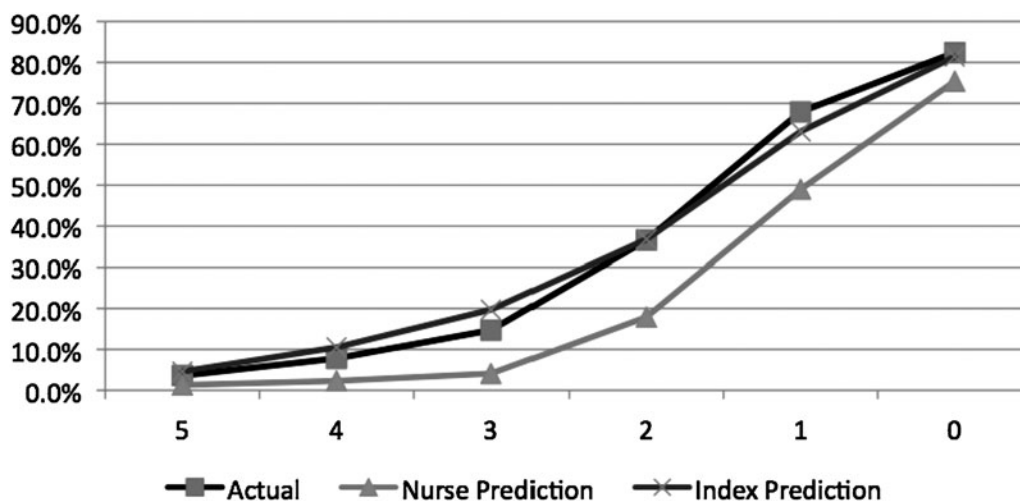
FIG. 1. Seven-day mortality of hospice patients in the validation sample ($n = 12,578$), by index score (0–5).

specialized settings such as the intensive care unit (ICU), where predictions may be more accurate.[27,28] It is noteworthy, though, that the hospice nurses in this study were subject to the same optimism bias that has been described elsewhere for physicians[24,25] and that they consistently underestimated 7-day mortality.

Second, this study found that a prognostic model predicts 7-day mortality more accurately than nurses do. This is surprising because the nurses in this sample had access to a wealth of clinical information (e.g., blood pressure, respiratory rate, overall appearance) that was not included in the model. More importantly, the survival index appears to offer a useful stratification tool that could help hospices and providers to be able to identify those patients at highest risk of dying within days of hospice admission, so that the necessary services can be provided quickly. In addition, the variables included in the model described here are all available at the time of hospice admission. This will make it possible for EHR-based decision support systems to calculate a patient's risk of 7-day mortality at the time of admission, creating a flag for hospice clinicians and referring physicians.

This study has three limitations that should be noted. First, although the development and validation samples described here were distinct, the hospice programs from which they were drawn are somewhat different than other hospices in the United States. For instance, they are larger than most,[2] they use an EHR system, and they have a low prevalence of non-white patients. Therefore, further research is needed to determine how well this index predicts survival in a broader sample of hospices.

Second, nurses' predictions were based on their assessments that death was "imminent," not that a patient would die within a week. It is possible that if this question had been asked directly, nurses' responses might have more closely resembled actual mortality rates. However, the current wording of the question provided the best c-statistic for predictions of death at one week, and therefore this timeframe is a reasonable test of nurses' predictions.

Third, we did not have information about the nurses whose predictions are reported here. For instance, we were unable to determine whether some were more accurate than others, or whether accuracy was associated with nurses' characteristics (e.g., experience, training). Future research is needed to define variation in accuracy among nurses, and to identify characteristics that are associated with more accurate predictions.

As the hospice industry grows and evolves, it will be essential to meet the needs of the patients who are referred very late. There may be opportunities to promote earlier referrals to hospice,[29] but gains are likely to be modest. Therefore, tools such as the index described here are needed to help health care providers meet patients' and families' needs.

Author Disclosure Statement

References

1. Center for Medicare and Medicaid Services: Medicare Hospice Regulations. 42 Code of Federal Regulations, Part 418.22; 1996. Washington, DC: US Government Printing Office.
2. National Hospice and Palliative Care Organization Facts and Figures. Hospice Care in America. 2010. www.nhpco.org/files/public/Statistics_Research/Hospice_Facts_Figures_Oct-2010.pdf [Last accessed March 1, 2012.]

3. Rickerson E, Harrold J, Carroll J, Kapo J, Casarett D: Timing of hospice referral and families' perceptions of services: Are earlier hospice referrals better? J Am Geriatr Soc 2005; 53:819–823.

4. Teno JM, Shu J, Casarett DJ, Spence C, Rhodes R, Connor SR: Timing of referral to hospice and quality of care: Length of stay and bereaved family members' perceptions of the timing of hospice referral. J Pain Symptom Manage 2007;30:120–125.

5. Younis T, Milch R, Abul-Khoudoud N, Lawrence D, Mirand A, Levine E: Length of survival in hospice for cancer patients referred from a comprehensive cancer center. Am J Hospice Palliat Med 2009;26:281–287.

6. Weng LC, Huang HL, Wilkie DJ, Hoenig NA, Suarez ML, Marschke M, Durham J: Predicting survival with the Palliative Performance Scale in a minority-serving hospice and palliative care program. J Pain Symptom Manage 2009;37:642–648.

7. Harrold J, Rickerson E, McGrath J, Morales K, Kapo J, Casarett D: Is the Palliative Performance Scale a useful predictor of mortality in a heterogeneous hospice population? J Palliat Med 2005;8:503–509.

8. Morita T, Tsunoda J, Inoue S, Chihara S: Validity of the palliative performance scale from a survival perspective. J Pain Symptom Manage 1999;18:2–3.

9. Altman DG, Royston P.:The cost of dichotomising continuous variables. BMJ 2006;332:1080.

10. Ambler G, Omar RZ, Royston P: A comparison of imputation techniques for handling missing predictor values in a risk model with a binary outcome. Stat Methods Med Res 2007;16:277–298.

11. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR: Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.

12. Marshall A, Altman DG, Royston P, Holder RL: Comparison of techniques for handling missing covariate data within prognostic modelling studies: A simulation study. BMC Med Res Methodol 2010;10:7.

13. Vergouwe Y, Royston P, Moons KGM, Altman DG: Development and validation of a prediction model with missing predictor data: A practical approach. J Clin Epidemiol 2010; 63:205–214.

14. Altman DG, Vergouwe Y, Royston P, Moons KGM: Prognosis and prognostic research: Validating a prognostic model. BMJ 2009;338:b605.

15. Royston P, Moons KGM, Altman DG, Vergouwe Y: Prognosis and prognostic research: Developing a prognostic model. BMJ 2009;338:b604.

16. Altman DG, Royston P: What do we mean by validating a prognostic model? Stat Med 2000;19:453–473.

17. Sun GW, Shook TL, Kay GL: Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis.J Clin Epidemiol 1996;49:907–916.

18. Steyerberg EW, Eijkemans MJ, Harrell FEJ, Habbema JD: Prognostic modeling with logistic regression analysis: In search of a sensible strategy in small data sets. Med Decis Making 2001;21:45–56.

19. Mantel N: Why stepdown procedures in variable selection? Technometrics 1970;12:621–625.

20. Steyerberg EW, Harrell FEJ, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD: Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774–781.

21. Akaike H: Information theory and an extension of the maximum likelihood principle. In: Petrov B, Csaki F (eds): *The Second International Symposium on Information Theory.* Budapest: Akademia Kiado, 1973, pp. 267–281.

22. Mallett S, Royston P, Dutton S, Waters R, Altman DG: Reporting methods in studies developing prognostic models in cancer: A review. BMC Med 2010;8:20.

23. Harrell FEJ, Lee KL, Califf RM, Pryor DB, Rosati RA: Regression modelling strategies for improved prognostic prediction. Stat Med 1984;3:143–152.

24. Christakis NA, Iwashyna T: Attitude and self-reported practice regarding prognostication in a national sample of internists. Arch Int Med 1998;158:2389–2395.

25. Christakis NA, Lamont EB: Extent and determinants of error in doctors' prognoses in terminally ill patients: Prospective cohort study. BMJ 2000;320:469–472.

26. Fromme EK, Smith MD, Bascom PB, Kenworthy-Heinige T, Lyons KS, Tolle SW: Incorporating routine survival prediction in a U.S. hospital-based palliative care service. J Palliat Med 2010;13:1439–1444.

27. Rocker G, Cook D, Sjokvist P, Weaver B, Finfer S, McDonald E, Marshall J, Kirby A, Levy M, Dodek P, Heyland D, Guyatt G: Clinician predictions of intensive care unit mortality. Crit Care Med 2004;32:1149–1154.

28. Copeland-Fields L, Griffin T, Jenkins T, Buckley M, Wise LC: Comparison of outcome predictions made by physicians, by nurses, and by using the Mortality Prediction Model. Am J Crit Care 2001;10:313–319.

29. Casarett D, Karlawish J, Crowley R, Mirsch T, Morales K, Asch DA: Improving use of hospice care in the nursing home: A randomized controlled trial. JAMA 2005;294:211–217.

Address correspondence to:
David J. Casarett, M.D., M.A.
Perelman School of Medicine
University of Pennsylvania
3615 Chestnut Street
Philadelphia, PA 19104

E-mail: casarett@mail.med.upenn.edu