Derivation and Validation of an In-Hospital Mortality Prediction Model Suitable for Profiling Hospital Performance in Heart Failure

Tara Lagu, MD, MPH; Penelope S. Pekow, PhD; Mihaela S. Stefan, MD; Meng-Shiou Shieh, PhD; Quinn R. Pack, MD, MSc; Mohammad Amin Kashef, MD; Auras R. Atreya, MD, MPH; Gregory Valania, DO; Mara T. Slawsky, MD, PhD; Peter K. Lindenauer, MD, MSc

Background—Comparing heart failure (HF) outcomes across hospitals requires adequate risk adjustment. We aimed to develop and validate a model that can be used to compare quality of HF care across hospitals.

Methods and Results—We included patients with HF aged ≥18 years admitted to one of 433 hospitals that participated in the Premier Inc Data Warehouse. This model (Premier) contained patient demographics, comorbidities, and acute conditions present on admission, derived from administrative and billing records. In a separate data set derived from electronic health records, we validated the Premier model by comparing hospital risk-standardized mortality rates calculated with the Premier model to those calculated with a validated clinical model containing laboratory data (LAPS [Laboratory-Based Acute Physiology Score]). Among the 200 832 admissions in the Premier Inc Data Warehouse, inpatient mortality was 4.0%. The model showed acceptable discrimination in the warehouse data (C statistic 0.75; 95% confidence interval, 0.74–0.76). In the validation data set, both the Premier model and the LAPS models showed acceptable discrimination (C statistic: Premier: 0.76 [95% confidence interval, 0.74–0.77]; LAPS: 0.78 [95% confidence interval, 0.76–0.80]). Risk-standardized mortality rates for both models ranged from 2% to 7%. A linear regression equation describing the association between Premier- and LAPS-specific mortality rates revealed a regression line with a slope of 0.71 (SE: 0.07). The correlation coefficient of the standardized mortality rates from the 2 models was 0.82.

Conclusions—Compared with a validated model derived from clinical data, an HF mortality model derived from administrative data showed highly correlated risk-standardized mortality rate estimates, suggesting it could be used to identify high- and low-performing hospitals for HF care. (*J Am Heart Assoc*. 2018;7:e005256. DOI: 10.1161/JAHA.116.005256.)

Key Words: heart failure • mortality • quality of care

H eart failure (HF) accounts for ≈1 million hospital admissions and $39 billion spent per year.[1,2] This large volume and high cost has led inpatient HF care to become a major focus of hospital quality measurement and improvement efforts, including public reporting of hospital-level mortality and readmission rates on websites such as Hospital Compare.[3]

From the Institute for Healthcare Delivery and Population Science (T.L., P.S.P., M.S.Stefan, M.S.Shieh, Q.R.P., P.K.L.), University of Massachusetts Medical School-Baystate (T.L., M.S.Stefan., Q.R.P., M.A.K., A.R.A., G.V., M.T.S., P.K.L.), Springfield, MA; Department of Medicine (T.L., M.S.Stefan., P.K.L.), Division of Cardiology (Q.R.P., M.A.K., A.R.A., G.V., M.T.S.), Baystate Medical Center, Springfield, MA; School of Public Health and Health Sciences, University of Massachusetts-Amherst, Amherst, MA (P.S.P.); Frankel Cardiovascular Center, University of Michigan, Ann Arbor, MI (A.R.A.).

Correspondence to: Tara Lagu, MD, MPH, 3601 Main St., 3rd Floor, IHDPS, Springfield, MA 01199. E-mail: tara.lagu@baystatehealth.org

Received December 7, 2016; accepted December 20, 2017.

To ensure that observed differences in outcomes between hospitals are not largely the result of differences in patient characteristics, care must be taken to adjust for case mix when describing differences in quality across hospitals. When estimating risk-standardized outcomes, analyses should also take into account hospital size and the clustered nature of the data (patients within hospitals) and should include only variables that are not related to the quality of care delivered.[4] Finally, the risk-adjustment model should be validated in different populations of patients and across hospitals. Although electronic health record (EHR) data are likely to become the data used for these purposes in the future, EHR data are not routinely available, meaning that there is still a need for risk-adjustment methods that do not use clinical data. The aim of this study was to develop and validate a model that can be used to compare quality of HF care across hospitals in situations in which clinical data are not available.

A recent innovation that will facilitate severity adjustment in claims data is the development of multihospital databases (eg, Premier, University HealthSystem Consortium) that

Clinical Perspective

What Is New?

- In a data set derived from electronic health records, we compared 2 models that calculate hospital mortality rates for patients with heart failure and found that a model that used only billing data performed very similarly to a model that used clinical data derived from medical records.

What Are the Clinical Implications?

- When using data sets that contain only information from hospitalizations (ie, those that lack prior outpatient data) or when clinical data are not available, our model that uses billing data could be useful for comparing quality of heart failure care across hospitals; it also allows for identification of hospitals with low mortality rates, creating an opportunity to conduct future studies that examine strategies of high-performing hospitals for heart failure care.

standardize highly detailed billing data across hospitals, providing time- or date-stamped information about all tests and services provided to individual patients.[5–9] Using one of these data sets, we sought to develop and validate a model that could be used to compare hospitals' performance in the care of HF patients. Then, in a separate hospital data set derived from hospital EHRs, we aimed to validate this model at the hospital level by comparing hospital risk–standardized mortality rates (RSMRs) for our model with RSMRs calculated from a validated model that uses laboratory results to predict mortality.

Methods

Derivation and Internal Validation Cohort

We gathered data from the cost-accounting systems of 433 hospitals that participated in the Premier Inc Data Warehouse (PDW; a voluntary, fee-supported database) between January 1, 2009, and June 30, 2011. PDW contains all elements found in hospital claims derived from the Uniform Billing 04 form. In addition, PDW contains an itemized, date-stamped log of all items and services charged to the patient or the insurer, including medications, diagnostic and therapeutic services, and laboratory tests. PDW includes ≈15% to 20% of all US hospitalizations. Participating hospitals are drawn from all regions of the United States, with greater representation from urban and southern hospitals. PDW has been used extensively for research purposes.[5–9] Because the data are proprietary, we are not able to make the data set or study materials available to other researchers.

We included patients who were aged ≥18 years and had a principal *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* diagnosis of HF or a principal diagnosis of respiratory failure with secondary diagnosis of HF when both HF and respiratory failure were coded "present on admission" (POA; *ICD-9-CM* codes for HF: 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 428.xx[10,11]; for respiratory failure: 518.81, 518.82, 518.84). Given the broad set of inclusion codes, we ensured that patients were treated for acute decompensated HF during the hospitalization by restricting the cohort to patients in whom at least 1 HF therapy (including diuretics, metolazone, inotropes, vasodilators, or intra-aortic balloon pump) was initiated within the first 2 days of hospitalization. In administrative data sets, the duration of the first hospital day includes partial days that can vary in length, so we chose the first 2 days of hospitalization (rather than just the first day) for initiation of an HF therapy. We excluded patients with a pediatric or psychiatric attending physician, those with elective admissions, and those who were transferred from or to another acute care facility (because we could not accurately determine the onset or subsequent course of their illness). For patients with repeat visits at a single hospital, 1 visit was randomly selected for inclusion. Patients were randomly assigned to a derivation cohort (80%) and an internal validation cohort (20%). The institutional review board at Baystate Medical Center granted permission to conduct the study and granted a waiver of informed consent because of the deidentified nature of the data.

External Validation Cohort

We validated the model in a population of patients with HF seen at hospitals that contributed to the HealthFacts database (Cerner Corp.) between January 2010 and December 2012. HealthFacts is a multihospital data set derived from the comprehensive EHRs of 116 geographically and structurally diverse hospitals throughout the United States. HealthFacts contains time-stamped pharmacy, laboratory, and billing information and contains records, including clinical data such as laboratory data, of >84 million acute admissions, emergency room visits, and ambulatory visits. We limited the sample to hospitals that contributed to the pharmacy, laboratory, and diagnosis segments of the database and had at least 20 eligible HF patients during the study period (to obtain stable hospital rates) after applying the same patient-level inclusion criteria.

Outcome

The primary outcome for both cohorts was hospital-specific risk-standardized all-cause in-hospital mortality.

Model Derivation

Patient predictors of mortality

Using the derivation cohort, we identified candidate variables that were used in prior risk adjustment models,[10,12,13] including patient age, sex, marital status, insurance status, and race/ethnicity. We used software provided by the Healthcare Costs and Utilization Project of the Agency for Healthcare Research and Quality[14,15] to identify the presence of comorbid conditions. In addition, we used POA codes to identify other acute conditions that are of concern in the setting of HF but that are not recognized in the Elixhauser comorbidity index. These conditions included atrial fibrillation, acute myocardial infarction, pneumonia, and acute kidney injury. Because we lacked echocardiogram results, we used *ICD-9-CM* codes to identify HF subtypes: systolic only, diastolic only, or both.

Premier model development

In the PDW population, we used a generalized estimating equation logistic regression model, clustering on hospital, to predict each patient's in-hospital mortality. We initially included all clinically relevant variables in the model, including variables with a well-established association with mortality (eg, age), all conditions included in the Elixhauser comorbidity index,[14] and selected comorbid acute illnesses (eg, acute myocardial infarction that was present on admission). Using backward selection, we retained variables in the final model (hereafter called the "Premier" model) with $P<0.05$. We calculated the area under the receiver operating characteristic curve (C statistic) and examined the model's calibration. We then applied the model coefficients to the validation cohort and examined model fit. Of note, we previously externally validated, at the patient level, the Premier model by comparing its performance to the performance of other published clinical HF models in a separate clinical data set.[16]

Laboratory-Based Acute Physiology Score model development

The Laboratory-Based Acute Physiology Score (LAPS) is a validated score that uses laboratory data derived from an EHR to predict in-hospital mortality across conditions, including HF. LAPS uses a 2-stage algorithm. First, selected variables are used to stratify patients into low and high mortality risk groups. Then 14 laboratory values (anion gap; albumin; arterial oxygen, pH, and carbon dioxide; bicarbonate; bilirubin; blood urea nitrogen; creatinine; glucose; hematocrit; sodium; troponin I; white blood cell count) are added to the algorithm to calculate a score.[17,18] For laboratory values that are not available, the algorithm assigns points based upon the patient's stage-1 mortality risk group (rather than using

imputation).[17,18] Because LAPS is designed to be used as a variable in a model that includes other patients characteristics when predicting mortality, we developed a generalized estimating equation logistic mortality prediction model using the LAPS score along with age, sex, race, and comorbidities. Using backward selection, we retained variables in the final model with $P<0.05$. We also previously externally validated (at the patient level, in a population with HF) the LAPS model by comparing its performance to the performance of other published clinical HF models.[16]

Calculation of RSMRs

For the external hospital-level validation of the Premier model in the HealthFacts data set, we compared RSMRs derived using the Premier model to those derived using the LAPS model. To calculate RSMRs for each hospital, we fit hierarchical generalized linear models using variables selected in the model-development step. We used generalized linear mixed models with a logit link to predict mortality and included covariates and a random effect for hospital. We assumed that random hospital effects are normally distributed and independent of hospital-level covariates. This method adjusts for within-hospital correlation of the observed outcomes. It also models the assumption that there are underlying differences between hospitals by allowing a random hospital intercept. For each model, we estimated RSMRs as the ratio of predicted mortality in each hospital given its patient mix and hospital-specific effect, divided by expected mortality given the patient mix and average hospital effect (ie, the mortality if the patients were treated at the "average" hospital). Next, we used bootstrap methods to develop a 95% confidence interval (CI) estimate of risk-standardized mortality for each hospital. To do this, we repeatedly sampled 55 hospitals with replacement (500 samples) from the 55 hospitals in the data set; we repeated the hierarchical generalized linear modeling with each sample and derived the risk-standardized mortality for each hospital, ultimately giving us ≈500 RSMR estimates for each hospital. We then used the 2.5 and 97.5 percentiles of the distribution for each hospital as the 95% CI estimate of RSMR. To compare RSMRs derived from the Premier and LAPS models, we first examined the distribution of RSMRs using histograms. We then used linear regression to model the association between the 2 rates, weighting each hospital by number of observations.[11] An intercept close to 0 and slope close to 1 would indicate similar RSMRs by the 2 models. We also calculated correlation coefficient between the RSMRs for Premier and LAPS, and the median difference between model estimates.

Description of cohort definition and sensitivity analyses

To better understand the HealthFacts validation cohort (which included both patients with a principal diagnosis of

HF and patients with a principal diagnosis of respiratory failure and secondary diagnosis of HF), we compared characteristics of patients with each principal diagnosis using χ^2 or Wilcoxon tests. We then conducted a sensitivity analysis limiting our cohort to patients with a principal diagnosis of HF and compared model fit to the results from the full cohort. All analyses were carried out using SAS version 9.4 (SAS Institute) and STATA version 13 (Stata-Corp).

Results

Derivation and Internal Validation Cohort

The PDW included 433 hospitals that contributed a total of 200 832 HF patients during the study period, with between 1 and 2125 patients per hospital (Table 1). Mean patient age was 73 years; approximately half were women (53%), and the majority (65%) were white. The principal diagnosis was HF in 90% of patients and respiratory failure in 10% of patients. Approximately 71% of patients had hypertension, 41% of patients had chronic obstructive pulmonary disease, 46% had diabetes mellitus, and 40% had chronic renal insufficiency. Patients commonly had additional POA acute diagnoses, including acute myocardial infarction (4%), acute kidney injury (17%), and atrial fibrillation (39%). Within the first 2 days of hospitalization, 19% of the patients were admitted to the intensive care unit, 5% received invasive mechanical ventilation, 6% received noninvasive ventilation, 5% received inotropes, and 7% received intravenous vasodilators. Overall, 8110 patients (4%) died in the hospital. The 80% derivation sample and the 20% internal validation sample were statistically similar to each other and to the larger cohort.

Validation Cohort

For external validation, we included 19 050 patients from 55 hospitals that contributed pharmacy, diagnosis, and laboratory data to the HealthFacts database. Compared with the PDW, the HealthFacts cohort had a higher percentage of black patients (28%), a slightly younger mean age (72), fewer Medicare patients (56% versus 77% in PDW), and a similar prevalence of comorbidities (Table 1). The rates of early mechanical ventilation (5%), noninvasive ventilation (7%), inotropes (6%), vasodilators (7%), and in-hospital death (4%) were also similar.

The majority (17 391 of 19 050, or 91.3%) of the included patients had a principal diagnosis of HF, whereas the remaining 1659 had a principal diagnosis of respiratory failure with a secondary diagnosis of HF (Table 2). Nearly all (18 737 of 19 050, or 98.3%) patients included in the

HealthFacts cohort received diuretics during their first 2 days of hospitalization. Among patients with a principal diagnosis of HF, 98.8% received diuretics within the first 2 days of hospitalization, and the remainder received ≥1 of the other HF therapies. Among patients with a principal diagnosis of respiratory failure, 93% received diuretics within the first 48 hours, and the remaining 7% received ≥1 of the other therapies (inotropes, vasodilators, or intra-aortic balloon pump).

There were further differences between patients with a principal diagnosis of HF and a principal diagnosis of respiratory failure. Patients with a principal diagnosis of respiratory failure and secondary diagnosis of HF were slightly younger than patients with a principal diagnosis of HF and were more likely to be white (versus other races), to have comorbid illnesses, or to have acute conditions (Table 2). The respiratory failure group also had a much higher mortality rate (13% versus 3%) than the group with a principal diagnosis of HF.

Model Performance

The Premier model showed acceptable calibration across deciles of predicted mortality ranging from 0.9% to 14.9% with C statistics of 0.75 (95% CI, 0.74–0.76) in the PDW cohort and 0.76 in the validation cohort (95% CI, 0.74–0.77). For further details, please see the prior publication[16]; for the model coefficients, see Table 3.

LAPS also showed good calibration in the validation cohort across deciles of predicted mortality ranging from 0.7% to 16.0% with a C statistic of 0.78 (95% CI, 0.76–0.80; for the model coefficients, see Table 4).

Premier Versus LAPS: Profiling Hospitals

RSMRs in the HealthFacts data had distributions that varied slightly by model. For both Premier and LAPS, distributions of hospital-level RSMRs ranged from 2% to 7% (Figure 1). The slope of the weighted regression line of the Premier- versus LAPS-specific mortality rates was 0.71 (SE: 0.07). The Pearson correlation coefficient of the standardized mortality rates from the 2 models was 0.82 ($P<0.001$; Figure 2). The median difference between RSMRs estimated from the models was 0.0001.

Sensitivity Analyses

We conducted a sensitivity analysis in which we applied the model developed in the Premier data set to a HealthFacts cohort limited to patients with a principal diagnosis of HF to determine if our inclusion of patients with a principal diagnosis of respiratory failure and secondary diagnosis

Table 1. Patient and Hospital Characteristics in Premier and HealthFacts

	Premier Data	HealthFacts Data	P Value
Total patients, N	200 832	19 050	
Number of hospitals	433	55	
Demographics			
Age, mean (median, Q1–Q3)	73.4 (76, 65–85)	72.0 (75, 62–84)	<0.0001
Female	105 413 (52.5)	9892 (51.9)	0.1380
Race/ethnicity			<0.0001
White	130 634 (65.0)	12 585 (66.1)	
Black	34 165 (17.0)	5304 (27.8)	
Hispanic	9442 (4.7)	341 (1.8)	
Other	26 591 (13.2)	820 (4.3)	
Insurance			<0.0001
Medicare	154 219 (76.8)	10 606 (55.7)	
Medicaid	13 133 (6.5)	1063 (5.6)	
Private	22 343 (11.1)	1906 (10.0)	
Uninsured	7257 (3.6)	629 (3.3)	
Other/unknown	3880 (1.9)	4846 (25.4)	
Clinical characteristics			
Comorbidities			
Valvular disease	4069 (2.0)	562 (3.0)	<0.0001
Pulmonary circulation disease	5423 (2.7)	559 (2.9)	0.0577
Peripheral vascular disease	26 472 (13.2)	1849 (9.7)	<0.0001
Hypertension	142 658 (71.0)	11 468 (60.2)	<0.0001
Paralysis	4010 (2.0)	322 (1.7)	0.0036
Other neurological disorders	16 712 (8.3)	1253 (6.6)	<0.0001
Chronic pulmonary disease	82 606 (41.1)	6588 (34.6)	<0.0001
Diabetes mellitus	91 594 (45.6)	7823 (41.1)	<0.0001
Hypothyroidism	36 652 (18.3)	2701 (14.2)	<0.0001
Renal failure	80 713 (40.2)	6976 (36.6)	<0.0001
Liver disease	5047 (2.5)	432 (2.3)	0.0379
Peptic ulcer disease, excluding bleeding	74 (0.0)	2 (0.0)	0.0615
AIDS	351 (0.2)	27 (0.1)	0.2928
Lymphoma	2012 (1.0)	161 (0.8)	0.0367
Metastatic cancer	2459 (1.2)	181 (1.0)	0.0009
Solid tumor without metastasis	3822 (1.9)	329 (1.7)	0.0879
Rheumatoid arthritis/collagen vascular disease	6193 (3.1)	511 (2.7)	0.0021
Coagulopthy	12 225 (6.1)	1061 (5.6)	0.0042
Obesity	38 337 (19.1)	3423 (18.0)	0.0002
Weight loss	10 218 (5.1)	750 (3.9)	<0.0001
Fluid and electrolyte disorders	67 083 (33.4)	5547 (29.1)	<0.0001
Chronic blood loss anemia	2331 (1.2)	134 (0.7)	<0.0001
Deficiency anemias	67 130 (33.4)	5194 (27.3)	<0.0001
Alcohol abuse	5997 (3.0)	491 (2.6)	0.0014
Drug abuse	3916 (1.9)	444 (2.3)	0.0003
Psychoses	6182 (3.1)	450 (2.4)	<0.0001
Depression	21 684 (10.8)	1461 (7.7)	<0.0001

Table 1. Continued

	Premier Data	HealthFacts Data	P Value
Comorbidity score mean (median, IQR)	4.8 (5, 3–6)	4.7 (4, 3–6)	<0.0001
≤2	30 966 (15.4)	4093 (21.5)	<0.0001
3–4	63 576 (31.7)	6157 (32.3)	
5–6	61 185 (30.5)	5451 (28.6)	
≥7	45 105 (22.5)	3349 (17.6)	
Acute and chronic conditions present on admission			
Acute myocardial infarction	8518 (4.2)	829 (4.4)	0.4706
Acute kidney injury	34 041 (17.0)	4105 (21.5)	<0.0001
Atrial fibrillation	78 434 (39.1)	6735 (35.4)	<0.0001
Coronary artery disease	108 125 (53.8)	9333 (49.0)	<0.0001
Pneumonia	30 097 (15.0)	2370 (12.4)	<0.0001
Malnutrition	6328 (3.2)	510 (2.7)	0.0003
Ischemic heart disease	107 818 (53.7)	9224 (48.4)	<0.0001
Heart failure type			
Systolic	78 012 (38.8)	8668 (45.5)	<0.0001
Diastolic	67 027 (33.4)	7438 (39.0)	<0.0001
Treatments			
Diuretics	198 244 (98.7)	18 737 (98.4)	<0.0001
Noninvasive ventilation	11 611 (5.8)	1324 (7.0)	<0.0001
Invasive mechanical ventilation	9969 (5.0)	939 (4.9)	0.8330
Vasopressors	3145 (1.6)	375 (2.0)	<0.0001
Inotropes	9586 (4.8)	1125 (5.9)	<0.0001
Vasodilators	15 024 (7.5)	1308 (6.9)	0.0020
Intra-aortic balloon pump	138 (0.1)	29 (0.2)	<0.0001
Prior year HF admissions			<0.0001
0	169 496 (84.4)	15 628 (82.0)	
1	22 634 (11.3)	2382 (12.5)	
≥2	8702 (4.3)	1040 (5.5)	
Outcomes			
Mortality	8110 (4.0)	800 (4.2)	0.2806
Hospital characteristics			
Urban (vs rural)	178 393 (88.8)	19 028 (99.9)	<0.0001
Teaching	75 626 (37.7)	15 299 (80.3)	<0.0001
Geographical location			<0.0001
Midwest	39 329 (19.6)	2593 (13.6)	
Northeast	40 522 (20.2)	7384 (38.8)	
South	87 607 (43.6)	7326 (38.5)	
West	33 374 (16.6)	1747 (9.2)	
Number of beds			<0.0001
<199	32 998 (16.4)	2143 (11.2)	
200–499	111 854 (55.7)	11 421 (60.0)	
≥500	55 980 (27.9)	5486 (28.8)	

Data are shown as n (%), unless otherwise noted. HF indicates heart failure; IQR, interquartile range.

of HF affected our results. We found that the Premier model in a cohort of patients with a principal diagnosis of HF had a C statistic of 0.76 (95% CI, 0.75–0.76) and a similar calibration curve to the full cohort. The C statistic for the LAPS model was also very similar to the main cohort (0.76; 95% CI, 0.74–0.78) and had a similar

Table 2. Comparison of Patients With Principal Diagnosis of HF to Those With Principal Diagnosis of Respiratory Failure With Secondary Diagnosis of HF

	Principal Diagnosis HF	Principal Diagnosis Respiratory Failure	P Value
Total patients	17 391 (91.3)	1659 (8.7)	
Demographics			
Age, mean (median, Q1–Q3)	72.2 (75, 62–85)	70.4 (71, 61–81)	<0.0001
Female	8922 (51.3)	970 (58.5)	<0.0001
Race/ethnicity			<0.0001
White	11 360 (65.3)	1225 (73.8)	
Black	4952 (28.5)	352 (21.2)	
Hispanic	310 (1.8)	31 (1.9)	
Other	769 (4.4)	51 (3.1)	
Insurance			<0.0001
Medicare	9561 (55.0)	1045 (63.0)	
Medicaid	953 (5.5)	110 (6.6)	
Private	1750 (10.1)	156 (9.4)	
Uninsured	574 (3.3)	55 (3.3)	
Other/unknown	4553 (26.2)	293 (17.7)	
Clinical characteristics			
Comorbidities			
Valvular disease	259 (1.5)	303 (18.3)	<0.0001
Pulmonary circulation disease	164 (0.9)	395 (23.8)	<0.0001
Peripheral vascular disease	1701 (9.8)	148 (8.9)	0.2583
Hypertension	10 356 (59.5)	1112 (67.0)	<0.0001
Paralysis	273 (1.6)	49 (3.0)	<0.0001
Other neurological disorders	1049 (6.0)	204 (12.3)	<0.0001
Chronic pulmonary disease	5466 (31.4)	1122 (67.6)	<0.0001
Diabetes mellitus	7056 (40.6)	767 (46.2)	<0.0001
Hypothyroidism	2419 (13.9)	282 (17.0)	<0.0001
Renal failure	6483 (37.3)	493 (29.7)	<0.0001
Liver disease	403 (2.3)	29 (1.7)	0.1367
Peptic ulcer disease x bleeding	2 (0.0)	0 (0.0)	0.6622
AIDS	26 (0.1)	1 (0.1)	0.3560
Lymphoma	145 (0.8)	16 (1.0)	0.5785
Metastatic cancer	147 (0.8)	34 (2.0)	<0.0001
Solid tumor without metastasis	279 (1.6)	50 (3.0)	<0.0001
Rheumatoid arthritis/collagen vascular disease	457 (2.6)	54 (3.3)	0.1309
Coagulopthy	910 (5.2)	151 (9.1)	<0.0001
Obesity	2944 (16.9)	479 (28.9)	<0.0001
Weight loss	611 (3.5)	139 (8.4)	<0.0001
Fluid and electrolyte disorders	4721 (27.1)	826 (49.8)	<0.0001
Chronic blood loss anemia	125 (0.7)	9 (0.5)	0.4118
Deficiency anemias	4697 (27.0)	497 (30.0)	0.0100
Alcohol abuse	433 (2.5)	58 (3.5)	0.0135
Drug abuse	400 (2.3)	44 (2.7)	0.3637
Psychoses	369 (2.1)	81 (4.9)	<0.0001
Depression	1278 (7.3)	183 (11.0)	<0.0001

Table 2. Continued

	Principal Diagnosis HF	Principal Diagnosis Respiratory Failure	P Value
Comorbidity score mean (median, IQR)	4.4 (4, 3–6)	5.1 (5, 3–6)	<0.0001
≤2	3951 (22.7)	142 (8.6)	<0.0001
3–4	5538 (31.8)	619 (37.3)	
5–6	4937 (28.4)	514 (31.0)	
≥7	2965 (17.0)	384 (23.1)	
Acute and chronic conditions present on admission			
Acute myocardial infarction	663 (3.8)	166 (10.0)	<0.0001
Acute kidney injury	3600 (20.7)	505 (30.4)	<0.0001
Atrial fibrillation	6174 (35.5)	561 (33.8)	0.1700
Coronary artery disease	8575 (49.3)	758 (45.7)	0.0049
Pneumonia	1789 (10.3)	581 (35.0)	<0.0001
Malnutrition	406 (2.3)	104 (6.3)	<0.0001
Ischemic heart disease	8486 (48.8)	738 (44.5)	0.0008
Heart failure type			
Systolic	8218 (47.3)	450 (27.1)	<0.0001
Diastolic	6867 (39.5)	571 (34.4)	<0.0001
Treatments			
Diuretics	17 191 (98.8)	1546 (93.2)	<0.0001
Noninvasive ventilation	921 (5.3)	403 (24.3)	<0.0001
Invasive mechanical ventilation	235 (1.4)	704 (42.4)	<0.0001
Vasopressors	200 (1.2)	175 (10.5)	<0.0001
Inotropes	907 (5.2)	218 (13.1)	<0.0001
Vasodilators	1112 (6.4)	196 (11.8)	<0.0001
Intra-aortic balloon pump	29 (0.2)	0 (0.0)	0.0960
Prior year HF admissions			0.7585
0	14 261 (82)	1367 (82.4)	
1	2174 (12.5)	208 (12.5)	
≥2	956 (5.5)	84 (5.1)	
Outcomes			
Mortality	577 (3.3)	223 (13.4)	<0.0001
Hospital characteristics			
Urban (vs rural)	17 369 (99.9)	1659 (100)	0.1472
Teaching	13 961 (80.3)	1338 (80.7)	0.7145
Geographical location			<0.0001
Midwest	2369 (13.6)	224 (13.5)	
Northeast	6828 (39.3)	556 (33.5)	
South	6591 (37.9)	735 (44.3)	
West	1603 (9.2)	144 (8.7)	
Number of beds			<0.0001
<199	1970 (11.3)	173 (10.4)	
200–499	10 348 (59.5)	1073 (64.7)	
≥500	5073 (29.2)	413 (24.9)	

Data shown as n (%), unless otherwise noted. HF indicates heart failure; IQR, interquartile range.

calibration curve. Because limiting the cohort to a principal diagnosis of HF excluded a large proportion of the deaths, with more deaths excluded from some hospitals than others (due to variation in coding of respiratory failure across hospitals), we opted not to calculate RSMRs in this limited cohort.

Table 3. Premier Model Coefficients From Development Sample (Premier Database), Derived Using Generalized Estimating Equation Modeling

Covariate	Estimate	SE	P Value
Intercept	−4.1603	0.0949	<0.0001
Demographics			
Age, y			
<50	Ref		
50–59	0.1305	0.0924	0.1579
60–69	0.2940	0.0885	0.0009
70–74	0.3784	0.0948	<0.0001
75–79	0.6849	0.0920	<0.0001
80–84	0.7570	0.0912	<0.0001
85–88	0.8364	0.0927	<0.0001
≥89	0.9584	0.0920	<0.0001
Race/ethnicity			
White	Ref		
Black	−0.2282	0.0463	<0.0001
Hispanic	−0.1202	0.0740	0.1042
Other	0.0521	0.0433	0.2297
Sex			
Male	Ref		
Female	−0.1209	0.0272	<0.0001
Insurance			
Medicare	Ref		
Private	−0.1016	0.0527	0.0540
Uninsured	−0.2119	0.1084	0.0507
Other/unknown	0.3766	0.0936	<0.0001
Medicaid	0.0245	0.0705	0.7284
Clinical characteristics			
Comorbidities			
Valvular disease	0.8410	0.0579	<0.0001
Pulmonary circulation disease	1.0029	0.0544	<0.0001
Peripheral vascular disease	0.1364	0.0361	0.0002
Hypertension	−0.2556	0.0281	<0.0001
Paralysis	0.5022	0.0727	<0.0001
Other neurological disorders	0.3745	0.0395	<0.0001
Chronic pulmonary disease	0.1245	0.0269	<0.0001
Renal failure	0.1817	0.0286	<0.0001
Liver disease	0.4563	0.0708	<0.0001
Rheumatoid arthritis	0.1936	0.0693	0.0052
Coagulopathy	0.6321	0.0405	<0.0001
Drug abuse	−0.2767	0.1344	0.0396
Lymphoma	0.3998	0.1051	0.0001
Solid tumor without metastasis	0.4095	0.0758	<0.0001

Continued

Table 3. Continued

Covariate	Estimate	SE	P Value
Metastatic cancer	0.8896	0.0794	<0.0001
Acute myocardial infarction	1.2192	0.0413	<0.0001
Acute kidney injury	0.7337	0.0305	<0.0001
Atrial fibrillation	0.2415	0.0269	<0.0001
Coronary artery disease	0.5637	0.1781	0.0015
Ischemic heart disease	−0.7037	0.1782	<0.0001
Pneumonia	0.6165	0.0301	<0.0001
Malnutrition	0.6160	0.0526	<0.0001
Systolic heart failure	−0.2137	0.0289	<0.0001
Diastolic heart failure	−0.3350	0.0298	<0.0001
Prior heart failure hospitalization			
0	Ref		
1	0.1394	0.0398	0.0005
>1	0.1794	0.0625	0.0041

Ref indicates referent.

Discussion

Using a mortality prediction model that showed good performance in a multihospital billing data set composed of inpatients with HF, we calculated hospital RSMRs in a multihospital data set containing information from >50 hospitals' EHRs. When we compared the Premier model's RSMRs with RSMR estimates derived from a clinical model that uses laboratory data, we found that the 2 models produced estimates of hospitals' risk-standardized mortality that were similar, with a correlation coefficient of 0.82. This suggests that the Premier model could be a useful tool for describing the quality of HF care provided by hospitals in situations in which clinical data are not available.

Our model builds on an existing hospital-profiling model used for hospitalized HF patients. Krumholz et al developed and validated an administrative claims–based risk-adjustment model for HF with the purpose of characterizing hospital quality. This model, which is used by the Centers for Medicare and Medicaid Services (CMS) to compare mortality rates across hospitals for public reporting purposes, produces results that are very similar to medical records–based models.[10,11,19,20] There are several key differences between our model and the CMS model. First, the CMS model was developed to predict 30-day mortality rates, whereas ours predicts in-hospital-mortality. Second, unlike our model, the CMS model contains all claims in the year before hospitalization. These claims are not available in data sets that contain only information on hospitalizations, such as PDW. Consequently, our model may be more broadly applicable to

Table 4. LAPS Model Coefficients From the HealthFacts Cohort, Derived Using Generalized Estimating Equation Modeling

Covariate	Estimate	SE	P Value
Intercept	−4.6381	0.2280	<0.0001
Demographics			
Age, y			
<50	Ref		
50–59	0.1185	0.2278	0.6030
60–69	0.1010	0.2154	0.6393
70–74	0.0480	0.2296	0.8343
75–79	0.2655	0.2199	0.2272
80–84	0.3230	0.2181	0.1387
85–88	0.6662	0.2168	0.0021
≥89	0.6472	0.2149	0.0026
Race/ethnicity			
White	Ref		
Black	−0.4706	0.1200	<0.0001
Hispanic	−0.4442	0.3352	0.1850
Other	−0.3141	0.2036	0.1229
Comorbidities			
Pulmonary circulation disease	0.5229	0.1521	0.0006
Hypertension	−0.3618	0.0806	<0.0001
Paralysis	0.6055	0.2171	0.0053
Other neurological disorders	0.5606	0.1162	<0.0001
Renal failure	0.1591	0.0787	0.0433
Coagulopathy	0.8795	0.1112	<0.0001
Solid tumor without metastasis	0.5547	0.2121	0.0089
Metastatic cancer	1.2072	0.2275	<0.0001
LAPS	0.0324	0.0017	<0.0001

LAPS indicates Laboratory-Based Acute Physiology Score; Ref, referent.

databases that contain information about hospitalizations but lack outpatient data. The CMS model does not include patients with a principal diagnosis of respiratory failure, and that limits its use to patients with HF as a principal diagnosis. The CMS model was also developed before widespread use of POA codes, so it did not originally include POA acute diagnoses, and POA acute diagnoses have not been added to the CMS model in recent years. In contrast, our model attempts to adjust for presenting severity by including some acute diagnoses that were present at the time of admission (eg, pneumonia or acute myocardial infarction). Finally, our model includes all patients aged >18 years, whereas the CMS model includes only Medicare patients, the vast majority of whom are aged ≥65 years. Despite these differences, our C

statistics compare favorably to the CMS model, which had a C statistic of 0.70.[11]

The validation of a model that can be used to profile hospital quality of HF care has some important implications. Because it can be used with almost any multihospital data set, this model can be used by groups of hospitals that come together to improve quality ("quality improvement collaboratives") and will allow such collaboratives to identify hospitals that need the most focused efforts. Perhaps more important, our model can be used to identify high-performing hospitals (eg, Figure 2). It is possible that hospitals with lower mortality rates use quality-improvement and care-coordination strategies to standardize the care of patients with HF that may lead to better patient outcomes. Although we cannot use this current data set to identify the strategies that are most likely to improve care, future studies should consider using qualitative methods, such as those described by Bradley and Krumholz,[21] to do so. These interviews might include questions about hospital units that specialize in the care of HF patients (including nurses, physicians, and managers with expertise in HF care), use of protocols for early identification of volume status, coordinated efforts to identify patients who would most benefit from intensive care and procedures, appropriate use of palliative care and hospice for patients with end-stage disease, and interventions to integrate team members.

This study has several strengths. First, we validated the model for use at the hospital level in a separate population with different demographic characteristics. Second, we previously demonstrated the patient-level performance of the Premier model compared with published clinical HF mortality prediction models.[16] Third, our model is one of only 2 clinical prediction models in the HF literature aimed specifically at benchmarking hospitals and is potentially more broadly applicable than the CMS model for the reasons stated. Among published HF models that predict mortality in hospitalized HF patients,[22–26] most contain clinical variables, such as systolic blood pressure and serum creatinine level.[22,23,25] In contrast, our model does not require clinical data but performed similarly to a model that does.

Despite these strengths, this study also has several limitations. First, our model predicts in-hospital mortality rather than 30- or 90-day mortality because our data sets do not contain information about postdischarge deaths. Although inpatient mortality is widely used in outcomes research,[26–28] there are concerns that hospitals with shorter length of stay or patterns of discharging end-of-life patients to palliative care or hospice might have lower in-hospital mortality but similar 30-day mortality.[29] This is also an issue because CMS quality metrics use 30-day rather than inpatient mortality. Second, the lack of clinical data also introduced an additional limitation: We could not determine HF etiology or ejection
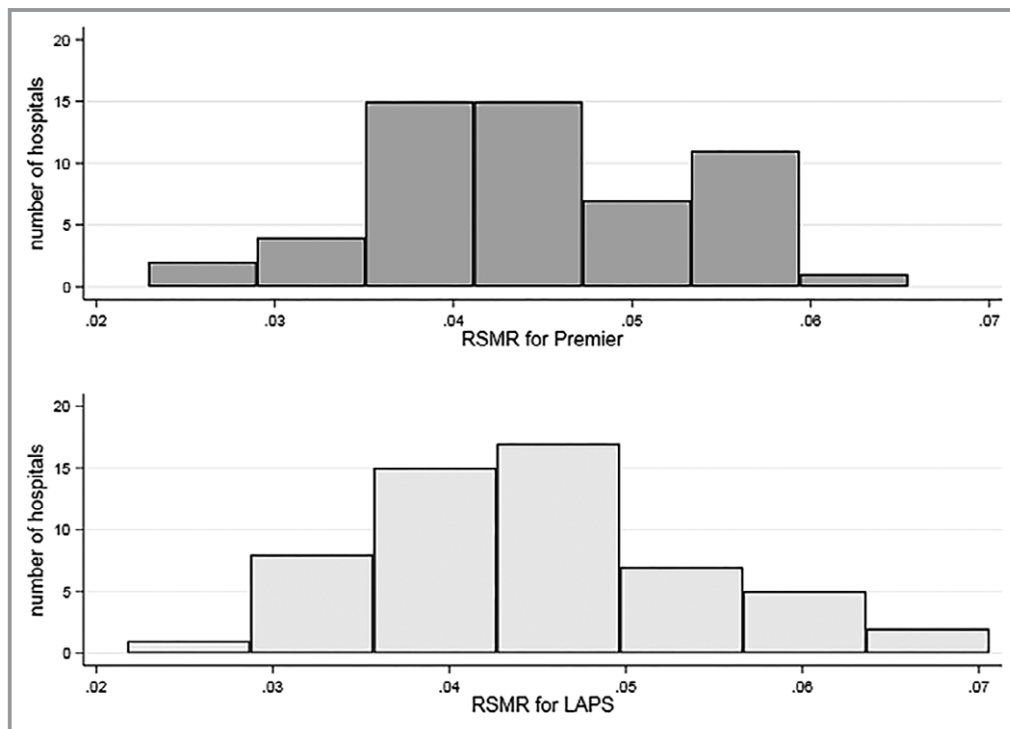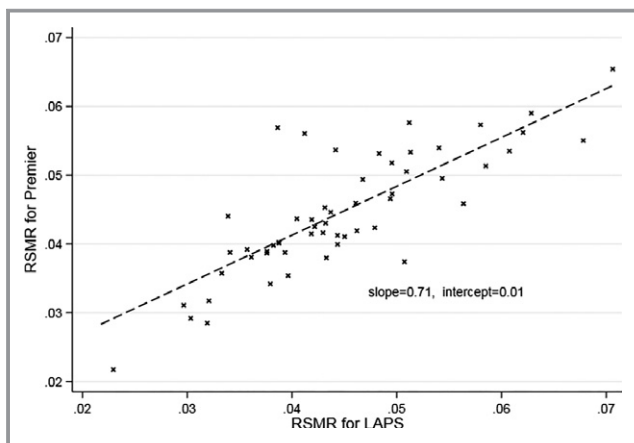
Figure 1. Distribution of hospital RSMRs for Premier and LAPS. LAPS indicates Laboratory-Based Acute Physiology Score; RSMR, risk-standardized mortality rate.

fraction, except by using diagnosis codes. Finally, our definition of HF was based on a combination of diagnosis and billing codes and does not include clinical data. We used a broad set of diagnosis codes to identify patients (with both HF and respiratory failure as principal diagnoses), but also required the presence of at least 1 HF-specific treatment early in the hospitalization as a proxy for clinical findings of HF. Although many prior studies and CMS quality metrics use diagnosis codes alone to identify patients hospitalized with HF,[10,30–32] we believe that using criteria that combine the principal diagnosis code with early therapies is a more specific method for identifying hospitalizations for which the primary issue is HF.[16] Furthermore, the high percentage of patients treated with diuretics provides additional evidence that these were hospitalizations for which the primary issue was HF. Slightly fewer patients with a principal diagnosis of respiratory failure received diuretics, but we feel that this is to be expected, given that these patients are the sickest and are most likely to suffer from low blood pressure, acute kidney injury, and other factors that would limit the use of diuretics but would increase the likelihood that they would receive other HF therapies (eg, inotropes). Finally, we used this broad set of diagnosis codes because prior work suggests that there is variation in coding across hospitals in the use of the code for acute respiratory failure, with some hospitals using this code more often than others.[33] If we failed to include patients with a principal diagnosis of respiratory failure for a hospital-level validation, the hospitals that use this code more frequently for their sicker patients with HF might appear to have a healthier cohort of patients with HF.

In conclusion, a mortality model designed to support hospital profiling of inpatient mortality for patients with HF appeared to produce similar results to a clinical model. These data suggest that this model could be useful for comparing quality of HF care across hospitals, especially if using data

Figure 2. Weighted regression line comparing LAPS-derived RSMRs with Premier-derived risk-standardized mortality. LAPS indicates Laboratory-Based Acute Physiology Score; RSMR, risk-standardized mortality rate.

sets that contain only information from hospitalizations (ie, those that lack prior outpatient data) or if clinical data are not available. This model also allows for identification of hospitals with low mortality rates, creating an opportunity to conduct future studies that examine strategies of high-performing hospitals for HF care.

Acknowledgments

Funding Sources

Disclosures

None.

References

1. Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, Ferguson TB, Ford E, Furie K, Gillespie C, Go A, Greenlund K, Haase N, Hailpern S, Ho PM, Howard V, Kissela B, Kittner S, Lackland D, Lisabeth L, Marelli A, McDermott MM, Meigs J, Mozaffarian D, Mussolino M, Nichol G, Roger VL, Rosamond W, Sacco R, Sorlie P, Roger VL, Stafford R, Thom T, Wasserthiel-Smoller S, Wong ND, Wylie-Rosett J. Heart disease and stroke statistics—2010 update: a report from the American Heart Association. *Circulation*. 2010;121: e46–e215.

2. Chen J, Normand S-LT, Wang Y, Krumholz HM. National and regional trends in heart failure hospitalization and mortality rates for Medicare beneficiaries, 1998–2008. *JAMA*. 2011;306:1669–1678.

3. Centers for Medicare and Medicaid Services. Hospital Compare. Available at: https://www.medicare.gov/hospitalcompare. Accessed January 23, 2018.

4. Krumholz HM, Brindis RG, Brush JE, Cohen DJ, Epstein AJ, Furie K, Howard G, Peterson ED, Rathore SS, Smith SC, Spertus JA, Wang Y, Normand S-LT; American Heart Association, Quality of Care and Outcomes Research Interdisciplinary Writing Group, Council on Epidemiology and Prevention, Stroke Council, American College of Cardiology Foundation. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. *Circulation*. 2006;113:456–462.

5. Lindenauer PK, Pekow PS, Lahti MC, Lee Y, Benjamin EM, Rothberg MB. Association of corticosteroid dose and route of administration with risk of treatment failure in acute exacerbation of chronic obstructive pulmonary disease. *JAMA*. 2010;303:2359–2367.

6. Twaddle ML, Maxwell TL, Cassel JB, Liao S, Coyne PJ, Usher BM, Amin A, Cuny J. Palliative care benchmarks from academic medical centers. *J Palliat Med*. 2007;10:86–98.

7. Reid M, Keniston A, Heller JC, Miller M, Medvedev S, Albert RK. Inappropriate prescribing of proton pump inhibitors in hospitalized patients. *J Hosp Med*. 2012;7:421–425.

8. Lindenauer PK, Pekow P, Wang K, Mamidi DK, Gutierrez B, Benjamin EM. Perioperative beta-blocker therapy and mortality after major noncardiac surgery. *N Engl J Med*. 2005;353:349–361.

9. Lindenauer PK, Pekow P, Wang K, Gutierrez B, Benjamin EM. Lipid-lowering therapy and in-hospital mortality following major noncardiac surgery. *JAMA*. 2004;291:2092–2099.

10. Krumholz H, Normand S-L, Galusha D, Mattera J, Rich A, Wang Y, Wang Y. Risk-adjustment models for AMI and HF 30-day mortality: methodology. Prepared for the Centers for Medicare & Medicaid Services under subcontract #8908-03-02. Yale University; 2005. Available at http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1163010421830. Accessed January 3, 2018.

11. Krumholz HM, Wang Y, Mattera JA, Wang Y, Han LF, Ingber MJ, Roman S, Normand S-LT. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. *Circulation*. 2006;113:1693–1701.

12. Higgins TL, Teres D, Copes WS, Nathanson BH, Stark M, Kramer AA. Assessing contemporary intensive care unit outcome: an updated Mortality Probability Admission Model (MPM0-III). *Crit Care Med*. 2007;35:827–835.

13. Lagu T, Lindenauer PK, Rothberg MB, Nathanson BH, Pekow PS, Steingrub JS, Higgins TL. Development and validation of a model that uses enhanced administrative data to predict mortality in patients with sepsis. *Crit Care Med*. 2011;39:2425–2430.

14. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36:8–27.

15. Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived rom ICD-9-CCM administrative data. *Med Care*. 2002;40:675–685.

16. Lagu T, Pekow PS, Shieh M-S, Stefan M, Pack QR, Amin Kashef M, Atreya AR, Valania G, Slawsky MT, Lindenauer PK. Validation and comparison of seven mortality prediction models for hospitalized patients with acute decompensated heart failure. *Circ Heart Fail*. 2016;9:e002912.

17. Escobar GJ, Gardner MN, Greene JD, Draper D, Kipnis P. Risk-adjusting hospital mortality using a comprehensive electronic record in an integrated health care delivery system. *Med Care*. 2013;51:446–453.

18. Escobar GJ, Greene JD, Scheirer P, Gardner MN, Draper D, Kipnis P. Risk-adjusting hospital inpatient mortality using automated inpatient, outpatient, and laboratory databases. *Med Care*. 2008;46:232–239.

19. Krumholz HM, Merrill AR, Schone EM, Schreiner GC, Chen J, Bradley EH, Wang Y, Wang Y, Lin Z, Straube BM, Rapp MT, Normand S-LT, Drye EE. Patterns of hospital performance in acute myocardial infarction and heart failure 30-day mortality and readmission. *Circ Cardiovasc Qual Outcomes*. 2009;2:407–413.

20. Bernheim SM, Grady JN, Lin Z, Wang Y, Wang Y, Savage SV, Bhat KR, Ross JS, Desai MM, Merrill AR, Han LF, Rapp MT, Drye EE, Normand S-LT, Krumholz HM. National patterns of risk-standardized mortality and readmission for acute myocardial infarction and heart failure. Update on publicly reported outcomes measures based on the 2010 release. *Circ Cardiovasc Qual Outcomes*. 2010;3:459–467.

21. Bradley EH, Curry LA, Ramanadhan S, Rowe L, Nembhard IM, Krumholz HM. Research in action: using positive deviance to improve quality of health care. *Implement Sci*. 2009;4:25.

22. Fonarow GC, Adams KF, Abraham WT, Yancy CW, Boscardin WJ; ADHERE Scientific Advisory Committee, Study Group, and Investigators. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *JAMA*. 2005;293:572–580.

23. Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting mortality among patients hospitalized for heart failure: derivation and validation of a clinical model. *JAMA*. 2003;290:2581–2587.

24. Eapen ZJ, Liang L, Fonarow GC, Heidenreich PA, Curtis LH, Peterson ED, Hernandez AF. Validated, electronic health record deployable prediction models for assessing patient risk of 30-day rehospitalization and mortality in older heart failure patients. *JACC Heart Fail*. 2013;1:245–251.

25. Peterson PN, Rumsfeld JS, Liang L, Albert NM, Hernandez AF, Peterson ED, Fonarow GC, Masoudi FA; American Heart Association Get With the Guidelines-Heart Failure Program. A validated risk score for in-hospital mortality in patients with heart failure from the American Heart Association get with the guidelines program. *Circ Cardiovasc Qual Outcomes*. 2010;3:25–32.

26. Tabak YP, Johannes RS, Silber JH. Using automated clinical data for risk adjustment: development and validation of six disease-specific mortality predictive models for pay-for-performance. *Med Care*. 2007;45:789–805.

27. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. *JAMA*. 2006;296:2694–2702.

28. Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. *Med Care*. 1992;30:615–629.

29. Drye E, Normand S-LT, Wang Y, Ross JS, Schreiner GC, Han L, Rapp M, Krumholz HM. Comparison of hospital risk-standardized mortality rates using in-hospital and 30-day models: implications for hospital profiling. *Ann Intern Med*. 2012;156:19–26.

30. Krumholz HM, Hsieh A, Dreyer RP, Welsh J, Desai NR, Dharmarajan K. Trajectories of risk for specific readmission diagnoses after hospitalization for heart failure, acute myocardial infarction, or pneumonia. *PLoS One*. 2016;11: e0160492.

31. Xu X, Li S-X, Lin H, Normand S-LT, Kim N, Ott LS, Lagu T, Duan M, Kroch EA, Krumholz HM. "Phenotyping" hospital value of care for patients with heart failure. *Health Serv Res*. 2014;49:2000–2016.

32. Partovian C, Gleim SR, Mody PS, Li S-X, Wang H, Strait KM, Allen LA, Lagu T, Normand S-LT, Krumholz HM. Hospital patterns of use of positive inotropic agents in patients with heart failure. *J Am Coll Cardiol*. 2012;60:1402–1409.

33. Lindenauer PK, Lagu T, Shieh M-S, Pekow PS, Rothberg MB. Association of diagnostic coding with trends in hospitalizations and mortality of patients with pneumonia, 2003–2009. *JAMA*. 2012;307:1405–1413.