Utilizing electronic health data and machine learning for the prediction of 30-day unplanned readmission or all-cause mortality in heart failure

Ashley N. Beecy, MD,* Manasa Gummalla, BA,† Evan Sholle, MS,‡ Zhuoran Xu, MSc,†§
Yiye Zhang, MSc, PhD,¶ Kelly Michalak, BA,† Kristina Dolan, BA,† Yasin Hussain, MD,‖
Benjamin C. Lee, PhD,† Yongkang Zhang, PhD,¶ Parag Goyal, MSc, MD,*
Thomas R. Campion Jr., PhD,‡¶ Leslee J. Shaw, PhD,†§ Lohendran Baskaran, MBBS,†§**
Subhi J. Al'Aref, MD††

From the *Division of Cardiology, Department of Medicine, Weill Cornell Medicine, New York, New York,
†Dalio Institute of Cardiovascular Imaging, Weill Cornell Medicine, New York, New York, ‡Information
Technologies & Services, Weill Cornell Medicine, New York, New York, ¶Department of Population
Health Sciences, Weill Cornell Medicine, New York, New York, §Department of Radiology, New
York-Presbyterian Hospital and Weill Cornell Medicine, New York, New York, ‖Yale Cardiovascular
Medicine, New Haven, Connecticut, **Department of Cardiovascular Medicine, National Heart
Centre, Singapore, and ††Division of Cardiology, Department of Medicine, University of Arkansas for
Medical Sciences, Little Rock, Arkansas.

BACKGROUND Existing risk assessment tools for heart failure (HF) outcomes use structured databases with static, single-timepoint clinical data and have limited accuracy.

OBJECTIVE The purpose of this study was to develop a comprehensive approach for accurate prediction of 30-day unplanned readmission and all-cause mortality (ACM) that integrates clinical and physiological data available in the electronic health record system.

METHODS Three predictive models for 30-day unplanned readmissions or ACM were created using an extreme gradient boosting approach: (1) index admission model; (2) index discharge model; and (3) feature-aggregated model. Performance was assessed by the area under the curve (AUC) metric and compared with that of the HOSPITAL score, a widely used predictive model for hospital readmission.

RESULTS A total of 3774 patients with a primary billing diagnosis of HF were included (614 experienced the primary outcome), with 796 variables used in the admission and discharge models, and

2032 in the feature-aggregated model. The index admission model had AUC = 0.723, the index discharge model had AUC = 0.754, and the feature-aggregated model had AUC = 0.756 for prediction of 30-day unplanned readmission or ACM. For comparison, the HOSPITAL score had AUC = 0.666 (admission model: P = .093; discharge model: P = .022; feature aggregated: P = .012).

CONCLUSION These models predict risk of HF hospitalizations and ACM in patients admitted with HF and emphasize the importance of incorporating large numbers of variables in machine learning models to identify predictors for future investigation.

KEYWORDS Big data; Electronic health data; Heart failure; Machine learning; Readmission

(Cardiovascular Digital Health Journal 2020;1:71–79) © 2020 The Authors. Published by Elsevier Inc. on behalf of Heart Rhythm Society. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Heart failure (HF) is a chronic medical condition that creates substantial economic burden for health care systems and is among the most common causes of hospital admission in the United States.[1–3] Despite advances in medical and percutaneous therapies, HF patients are vulnerable to frequent hospital admissions, with a mean estimated cost of

$23,000 per hospitalization and total annual costs predicted to rise to $70 billion by 2030.[4–7] Multiple inpatient hospitalizations are especially common after an initial HF diagnosis, with 1 survey showing that up to 42% of patients were hospitalized at least 4 times over mean follow-up of 4.7 years.[8] The rising financial burden associated with HF hospitalizations led to public reporting by the Centers for Medicare and Medicaid Services of risk-adjusted unplanned all-cause readmission rates among patients with HF.[9] As such, the 30-day readmission rate for HF patients has become a key quality metric, and subsequent

KEY FINDINGS

- Machine learning (ML) can be leveraged to build a risk prediction model for the composite of hospital readmission and all-cause mortality at 30 days in a heart failure population. Feature-aggregated models that include extensive social, clinical, and physiological parameters available from invasive and noninvasive studies can outperform the HOSPITAL score, which is a current clinical tool used for readmission risk prediction.

- ML models can deliver new insight into complex interactions, nonlinearities, and the unknown importance of prominently featured variables. Identifying prominently featured variables in the model, whether previously identified or novel, can form the basis for future investigations.

- Advanced analytics are needed, especially in heart failure and other complex diseases. Future studies are needed to determine the feasibility of incorporating disease-specific predictive models into the electronic health record and whether they improve patient care or lead to more efficient and cost-effective health care resource utilization.

reimbursement policies have sought to incentivize health care facilities to lower 30-day readmissions by reducing Medicare reimbursement for hospitals with high readmission rates.[10]

Given the enormous clinical and economic burden associated with HF management, mechanisms capable of predicting patient-level risk for readmission and/or mortality in the HF population have been the focus of significant research efforts. Numerous readmission risk prediction models have been developed, the best of which have displayed only modest predictive ability.[11–18] Most recently, machine learning (ML) has been combined with a big data approach for the prediction of all-cause 30-day readmissions producing minimal improvement over conventional statistics-based methods.[19,20] We hypothesize that previous approaches failed to yield higher predictive value because they utilized structured databases that incorporate limited patient characteristics that fail to capture the complex underlying pathogenesis of HF. To address this gap, we sought to build a higher-performing predictive model for 30-day unplanned readmissions or all-cause mortality in HF patients, using electronic health record (EHR) data that include extensive social, clinical, and physiological parameters available from invasive and noninvasive studies.

Methods

The CLEVER-HEART (PrediCtion Of EarLy REadmissions In Patients With CongestiVE HeaRt Failure: A Novel Approach) study aims to develop a predictive model that accurately describes the risk of 30-day *unplanned* readmissions or all-cause mortality in patients with HF, based on

an integrated, big data approach that includes psychosocial factors, demographics, and clinical data, and incorporates invasive and noninvasive hemodynamic and physiological parameters. The study was conducted at New York-Presbyterian Hospital/Weill Cornell Medicine (NYP-WCM), a quaternary care center and inpatient facility with over 2500 beds and hundreds of thousands of patients annually. The Dalio Institute of Cardiovascular Imaging has made use of the Architecture for Research Computing in Healthcare program, a suite of tools and services offered by the Research Informatics team within the Information Technologies & Services Department of WCM in conjunction with the Clinical and Translational Science Center and the Joint Clinical Trials Office.

Study population

The study population included all patients admitted or readmitted to the general medical or cardiac services at NYP-WCM between January 2008 and September 2018 and assigned a billing diagnosis of acute HF or acute on chronic HF, as defined by an *International Classification of Diseases, Ninth Revision* (ICD-9) code of 428.* or an *International Statistical Classification of Diseases, Tenth Revision* (ICD-10) code of I50.* The index admission was defined as the first inpatient admission with a diagnosis of HF within the study timeframe. A readmission was considered any subsequent inpatient admission. In total, the study cohort comprised 3774 patients assigned a primary billing code of HF at index discharge. We excluded patients who died during the course of their first HF hospitalization, were discharged to hospice from their first HF hospitalization, or were lost to follow-up within 30 days of discharge from the first HF hospitalization (Figure 1). Patients were considered lost to follow-up if they did not have any office visit or laboratory or imaging data recorded in the EHR within 30 days of index discharge. The WCM institutional review board approved the study protocol.

Data extraction and audit

Leveraging existing WCM infrastructure for the secondary use of the EHR for research, the study extracted an extensive
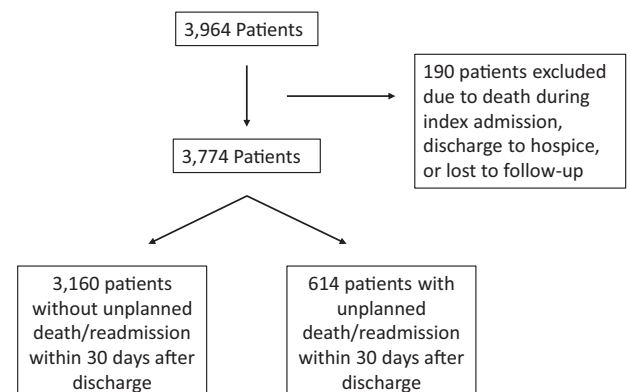
Figure 1 Flowchart of the included and excluded cohorts for the primary outcome.

dataset from NYP-WCM EHR systems, including but not limited to EpicCare-Ambulatory, Allscripts Sunrise Care Manager, Xcelera, and CoPath.[21] All data were extracted according to defined specifications and stored in Microsoft SQL Server architecture maintained by the Research Informatics team. Nine hundred eighty variables were retrieved, including socioeconomic, clinical history, laboratory, radiologic investigations, electrocardiogram (ECG), and invasive coronary angiographic measurements (Supplemental Table A.1). Patients with missing clinical information (n = 1606), not retrievable using automated extraction methods, were subject to manual chart review for extraction of the required clinical information. This information was recorded in REDCap (Research Electronic Data Capture), a Health Insurance Portability and Accountability Act–compliant research electronic data capture platform.[22]

To verify the accuracy of the data extraction, 200 patients were selected for an internal audit. Patients selected for audit were stratified by year of index admission and by readmission status (never readmitted, readmitted within 30 days, or readmitted after 30 days). For each of the selected audit cases, data extracted from the EHR using automated techniques were manually compared to the results of manual abstraction from the inpatient and outpatient EHR systems.

Primary and secondary outcomes

The primary outcome of the study was the occurrence of unplanned readmission or death from any cause within 30 days of discharge from the index HF hospitalization. A secondary outcome was the occurrence of 6-month unplanned readmission or all-cause death. An unplanned readmission was defined using the Centers for Medicare and Medicaid Services algorithm, a well-established technique for identifying hospital admissions attributed to unavoidable causes, such as scheduled organ transplantation or maintenance chemotherapy.[23] All-cause mortality data were derived from multiple sources, including inpatient and outpatient EHR data, as well as the Social Security Death Master File. Specifically, to ensure that nonevent patients were alive at the 1-year follow-up mark, (1) patients without follow-up data beyond the index admission were excluded from the database; and (2) the last follow-up date was considered the date an individual had EHR data.

Feature engineering and statistical analysis

To construct the model, we divided the variables into (1) static and (2) dynamic characteristics. Static characteristics constituted a set of static baseline level variables such as demographics, socioeconomic variables, and baseline clinical characteristics (eg, presence or absence of a given comorbidity). Dynamic characteristics, or time-dependent variables (eg, vital signs), consisted of multivariate time-series observations with varying lengths of sequences and irregular sampling. In model construction, we performed feature engineering before fitting models to transfer time-series classification problem into cross-sectional models. We categorized our predictive models into 3 groups: (1) index admission model; (2) index discharge model; and (3) feature-aggregated model. The index admission model is cross-sectional and comprised variables that were measured at the beginning of the index hospital stay. The index discharge model is similarly cross-sectional and comprised variables measured at the end of the index hospital stay. The feature-aggregated model incorporated longitudinal attributes and temporal information of time-dependent variables using summary statistics. Each time-dependent variable was aggregated into 7 descriptive statistics to represent the longitudinal attributes of the variables: minimum, maximum, standard deviation, mean, weighted average, average absolute change, and last report value within 6 months before discharge. The time-dependent variables were collected during the index hospital stay. For the index admission and discharge models, if a given variable during the index hospital stay was not available, we considered the last instance that was closest (within 6 months) to the admission and discharge dates, respectively.

Continuous variables are given as mean ± SD or median (interquartile range). Categorical variables are given as absolute value and proportion. Baseline characteristics of patients who experienced the primary outcome (all-cause death or unplanned readmission within 30 days) vs patients without the primary outcome were compared using the Student t test or Wilcoxon rank-sum test for continuous variables and the χ^2 or Fisher exact test for categorical variables, as appropriate. $P < .05$ was considered significant for all analyses. The prediction model for all-cause death or readmission within 30 day was constructed using an eXtreme Gradient Boosting (XGBoost) approach, which has been extensively validated as an accurate approach that provides the ability to utilize both continuous and categorical inputs while allowing for the handling of sparsity without the need for high computational power.[24] Imputation of missing values was not performed because XGBoost has intrinsic mathematical mechanisms to handle missing values. Furthermore, imputation in the context of EHR data might lead to biases because missingness is not present at random in such a context.

Models were constructed and tested using derivation, testing, and validation cohorts. The derivation and validation cohorts comprised 90% of the data; the remaining 10% of the data was used as an independent test cohort. The independent test data were not included in the data used to train the model. To develop derivation and validation cohorts, a stratified, 5-fold cross-validation was used. The study population was divided randomly into 5 subsets with similar event rates. To form the derivation cohort, 4 subsets were combined, and the remaining subset was reserved as the validation set. This process was repeated 5 times, such that every subset served as the validation set, thereby accounting for variability among patients and providing risk estimates for all cases (Figure 2). The area under the receiver operating curve (AUC) from the test set was used to determine the model performance and was compared to the HOSPITAL score
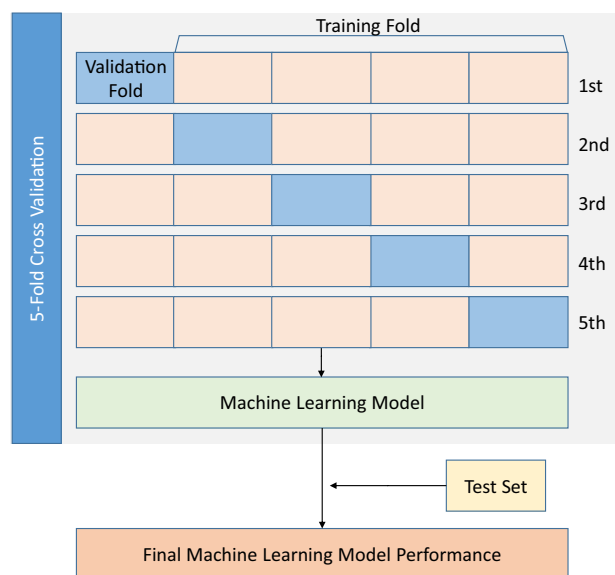
Figure 2   Analysis flow for the development and evaluation of models.

using the DeLong test,[25,26] which is a validated score that is commonly used in clinical practice for the identification of individuals at risk for avoidable 30-day readmission.[27] The HOSPITAL score incorporates the following variables: hemoglobin at discharge, discharge from oncology service, sodium level at discharge, any ICD coded procedure performed during hospital stay, index admission type, number of hospital admissions during the previous year, and length of stay. We also compared the performance of our model to that of serum brain natriuretic peptide (BNP) levels at admission.

Results
Baseline characteristics
The study population comprised 3774 primary HF subjects who met the inclusion and exclusion criteria. Six hundred fourteen (17.0%) experienced the primary outcome, including 92 (15.0%) instances of 30-day all-cause mortality and 522 (85.0%) unplanned 30-day readmissions. Table 1 lists the baseline characteristics of the studied cohort, stratified by primary outcome. Mean age of the entire cohort was 73.07 ± 15.23 years; 54.7% were male; and 51.1% were Caucasian. There was a high prevalence of cardiovascular risk factors (35.5% diabetes mellitus; 63.7% hypertension; 49.2% hyperlipidemia; 25.7% history of smoking). In addition, comorbid medical conditions were highly prevalent (30.9% chronic kidney disease; 10.9% chronic obstructive pulmonary disease; 50% coronary artery disease; 48.9% previous diagnosis of HF before index admission). In terms of New York Heart Association (NYHA) functional class, 17.1% were classified as class I, 28.7% as class II, 43.4% as class III, and 10.8% as class IV. The NYHA class was reported based on the patient's baseline functional status before the acute or acute on chronic HF decompensation leading to admission.

Mean hemoglobin concentration was 11.2 g/dL, and mean serum creatinine was 1.29 mg/dL. Mean BNP level was 862 pg/mL. On echocardiography, median left ventricular ejection fraction (calculated using the Teichholz formula) was 42.2% (interquartile range 25.1, 59.9). Mean left ventricular internal dimension in diastole was 5.6 cm. In individuals with measured values, mean pulmonary artery systolic pressure was elevated at 49.5 mm Hg (interquartile range 40.1, 61.1). In terms of admission medications, 81.4% were receiving beta-blocker therapy at the time of index admission, and 92.3% were receiving diuretic therapy. A small proportion of individuals (4.8%) was receiving inotropic therapy at the time of index admission.

Prediction of 30-day and 6-month outcomes
Three models were constructed for prediction of 30-day (primary outcome) or 6-month (secondary outcome) unplanned readmission or all-cause mortality, and compared prediction with that of the HOSPITAL score for avoidable readmissions. Seven hundred ninety-six variables were used in the admission and discharge models (184 variables from the original 980 had significant sparsity and were not used), and 2032 variables were used in the feature-aggregated model. Both the Goodman-Kruskal τ and χ^2 tests confirmed the presence of a significant correlation between variable sparsity and 30-day outcomes, and as such variable sparsity was included in the prediction model. The index admission model had AUC = 0.723, whereas the index discharge model had AUC = 0.754 for prediction of 30-day outcomes. In comparison, the HOSPITAL score had AUC = 0.666 (admission: $P = .093$; discharge: $P = .022$). The feature-aggregated model achieved AUC = 0.756 for prediction of 30-day outcomes, a value significantly higher than that of the HOSPITAL score ($P = .012$) and index admission BNP ($P < .01$) (Figure 3). In terms of the secondary outcome (236 patients were excluded from the secondary outcome because they had no follow-up data and were presumed to have been lost to follow-up, resulting in n = 3538), there was no significant difference between the 3 models and the HOSPITAL score for prediction of outcomes (admission: $P = .40$; discharge: $P = .20$; feature-aggregated: $P = .20$), whereas index admission BNP was significantly worse at predicting 6-month outcomes (AUC = 0.522; $P < .01$ for all comparisons).

Variable importance
Table 2 lists (in descending order) each variable's importance within the 3 models for both 30-day and 6-month unplanned readmissions or all-cause mortality. For 30-day outcomes, discharge to home (vs rehabilitation or long-term care facility), serum chemistry values (eg, hemoglobin level, red blood cell distribution width), and quantitative ECG variables (eg, R-wave axis, QRS duration, QTc interval, atrial rate) featured prominently for the 3 models. The echocardiographic variable with highest importance in the 30-day outcomes models was aortic valve area in the index discharge model.

Table 1   Baseline demographic and clinical data of the studied cohort stratified by primary outcome

Variable	Entire cohort (N = 3774)	No death/readmission within 30 days (n = 3160)	Death/readmission within 30 days (n = 614)	P value
Age (y)	73.07 ± 15.23	72.65 ± 15.17	75.25 ± 15.37	.0001
Male gender	2065 (54.72)	1761 (55.73)	304 (49.51)	.0046
Caucasian	1319 (51.08)	1105 (50.69)	214 (53.23)	.3481
Coronary artery disease	1885 (49.95)	1551 (49.08)	334 (54.4)	.0159
Myocardial infarction	982 (26.02)	783 (24.78)	199 (32.41)	.0001
Diabetes mellitus	1338 (35.45)	1109 (35.09)	229 (37.3)	.2967
Hypertension	2402 (63.65)	1983 (62.75)	419 (68.24)	.0097
Hyperlipidemia	1858 (49.23)	1537 (48.64)	321 (52.28)	.0987
History of smoking	968 (25.65)	781 (24.72)	187 (30.46)	.0029
Hemoglobin (g/dL)	11.2 (9.8, 12.7)	11.4 (9.9, 12.9)	10.7 (9.3, 12.11)	<.0001
Estimated GFR (mL/min/m^2)	42 (29, 53)	43 (29.75, 53)	40 (25.92, 52)	.0428
Serum potassium (mmol/L)	4.05 (3.75, 4.4)	4.05 (3.75, 4.4)	4.1 (3.8, 4.5)	.0097
Brain natriuretic peptide (pg/mL)	862 (412.5, 1640.25)	843 (401.5, 1587.75)	913 (459, 1896)	.0443
Beta-blocker	3001 (81.35)	2517 (81.46)	484 (80.8)	.7064
NYHA functional class				
I	635 (17.07)	545 (17.48)	90 (14.98)	.0058
II	1069 (28.74)	923 (29.6)	146 (24.29)	
III	1614 (43.4)	1320 (42.33)	294 (48.92)	
IV	401 (10.78)	330 (10.58)	71 (11.81)	

Values are given as mean ± SD, n (%), or median (interquartile range) unless otherwise indicated.
GFR = glomerular filtration rate; NYHA = New York Heart Association.

None of the invasive hemodynamic variables (including right-sided pressures) was ranked among the top 25 predictors. For 6-month outcomes, socioeconomic variables (eg, distance from home to a park, total population) became more important.

To further elucidate the association between the highest-ranked variables in the prediction models and their influence on 30-day and 6-month outcomes, multivariate logistic regression models were constructed for each of the 3 models for both the primary and secondary outcomes. Figure 4 shows the odds ratios and associated 95% confidence intervals for the 10 most important variables in each model. Lower hemoglobin and/or hematocrit were associated with lower occurrence of outcomes, whereas the presence of diuretic therapy in the discharge medication profile was significantly associated with readmission or all-cause mortality at 6 months.

Discussion

In the present analysis, we utilized an integrated approach using EHR data to develop 2 cross-sectional models and 1 feature-aggregated model of 30-day and 6-month unplanned readmissions or all-cause mortality. We found that incorporation of clinical, socioeconomic, ECG, invasive and noninvasive imaging, and hemodynamic parameters as well as administered medications resulted in improved outcome prediction in comparison to the HOSPITAL score, a commonly used risk score that predicts 30-day potentially avoidable
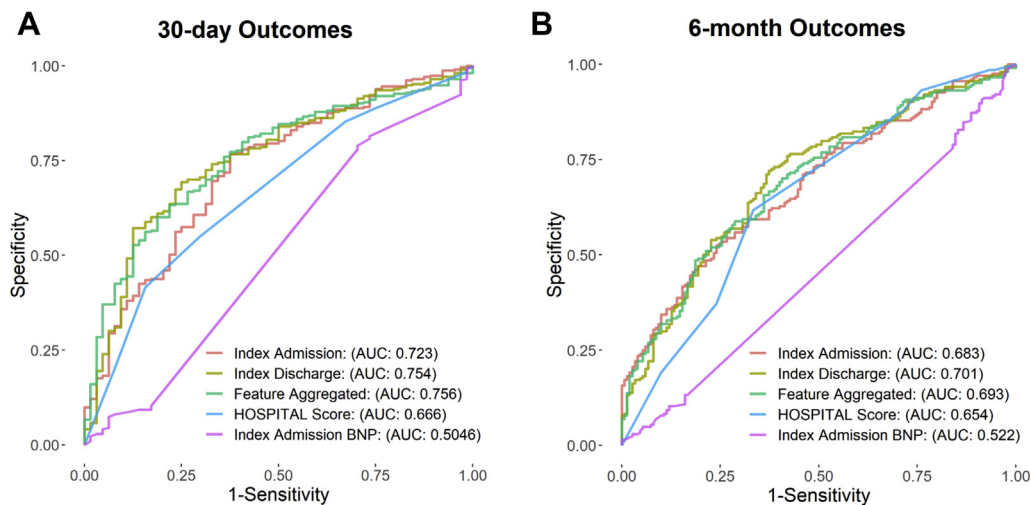
Figure 3   Area under the curve (AUC) plots for index admission model, index discharge model, and feature-aggregated model, and comparison to the HOSPITAL score and admission brain natriuretic peptide (BNP) levels for prediction of 30-day (**A**) and 6-month (**B**) unplanned readmission or all-cause mortality.

Table 2 Variable importance ranking for the 3 models for 30-day (primary outcome) or 6-month (secondary outcome) unplanned readmission or all-cause mortality

30-Day outcomes			6-Month outcomes		
Index admission	Index discharge	Feature aggregated	Index admission	Index discharge	Feature aggregated
Red cell distribution width (1)	Red cell distribution width (1)	Discharge to home	QRS duration (2)	QRS duration (2)	QRS duration mean (2)
Hemoglobin (1)	Hemoglobin (1)	Red cell distribution width max (1)	Serum creatinine (1)	Atrial rate (2)	Diuretic therapy (5)
R-wave axis (2)	R-wave axis (2)	Red cell distribution width mean (1)	Red cell distribution width (1)	Red cell distribution width (1)	History of psychiatric disease
Atrial rate (2)	Discharge to home	Hemoglobin min (1)	Diuretic (5)	Serum creatinine (1)	Antibiotic therapy (5)
Hematocrit (1)	WBC count (1)	Red cell distribution width last (1)	Blood urea nitrogen (1)	Noninvasive blood pressure (diastolic) (1)	Serum creatinine mean (1)
Blood urea nitrogen (1)	Atrial rate (2)	Hemoglobin mean (1)	Antibiotic therapy (5)	Park distance (4)	Serum magnesium (standard deviation) (1)
Lymphocyte absolute count (1)	Eosinophilia count (1)	R-wave axis mean (2)	History of psychiatric disease	Total population (4)	QTc interval min (2)
Index stay	T-wave axis (2)	T-wave axis weighted average (2)	T-wave axis (2)	Albumin (1)	Serum creatinine min (1)
QTc interval (2)	QRS duration (1)	PR interval mean (2)	Time to furosemide (6)	PR interval (2)	T-wave axis mean (2)
Platelet count (1)	Glucose (1)	Serum magnesium (standard deviation) (1)	Total population (4)	Mean corpuscular volume (1)	Red cell distribution width max (1)
QRS duration (2)	Index weight	R-wave axis weighted average (2)	Index stay length	P-wave axis (2)	Lymphocyte absolute count mean (1)
Glucose (1)	Potassium (1)	Atrial rate weighted average (2)	Atrial rate (2)	Activated PTT (1)	Total population (4)
Discharge to home	Ventricular rate (2)	Hematocrit min (1)	Hematocrit (1)	Lymphocyte absolute count (1)	Red cell distribution width mean (1)
QT interval (2)	Respiratory hazard index (4)	Hematocrit max (1)	PR interval (2)	Blood urea nitrogen (1)	Atrial rate mean (2)
Noninvasive blood pressure (1)	Albumin (1)	QTc interval min (2)	Lymphocyte absolute count (1)	Glucose (1)	QRS duration min (2)
Mean corpuscular volume (1)	QTc interval (2)	WBC count (standard deviation) (1)	Ventricular rate (2)	T-wave axis (2)	Cardiac medication (5)
Neutrophil absolute count (1)	Index stay	Cardiac medication (5)	Alkaline phosphatase (1)	Hemoglobin (1)	Noninvasive blood pressure (systolic) mean (1)
Ventricular rate (2)	Vitamin (6)	T-wave axis last (2)	History of permanent pacemaker	WBC count (1)	Hematocrit min (1)
Alkaline phosphatase (1)	Current weight (1)	Glucose (standard deviation) (1)	Serum BNP (1)	Noninvasive blood pressure (systolic) (1)	Anion gap (standard deviation) (1)
Serum calcium (1)	Noninvasive blood pressure (1)	Mean corpuscular volume (standard deviation) (1)	QT interval (2)	QT interval (2)	QTc interval (standard deviation) (2)
Total urine output	LA area (2) chamber view (3)	T-wave axis (standard deviation) (2)	Park distance (4)	Diuretic therapy (5)	Beta-blocker therapy (5)
Alanine aminotransferase (1)	Aortic valve area (3)	T-wave axis (average absolute change) (2)	Total urine output	Index stay	PR interval min (2)
Aspartate aminotransferase (1)	Cardiac medication (5)	Mean corpuscular volume max (1)	Hemoglobin (1)	History of psychiatric disease	Serum sodium mean (1)
T-wave axis (2)	P-wave axis (2)	Age	E/E' ratio (3)	Heart rate (1)	T-wave axis min (2)

Table 2   (Continued)

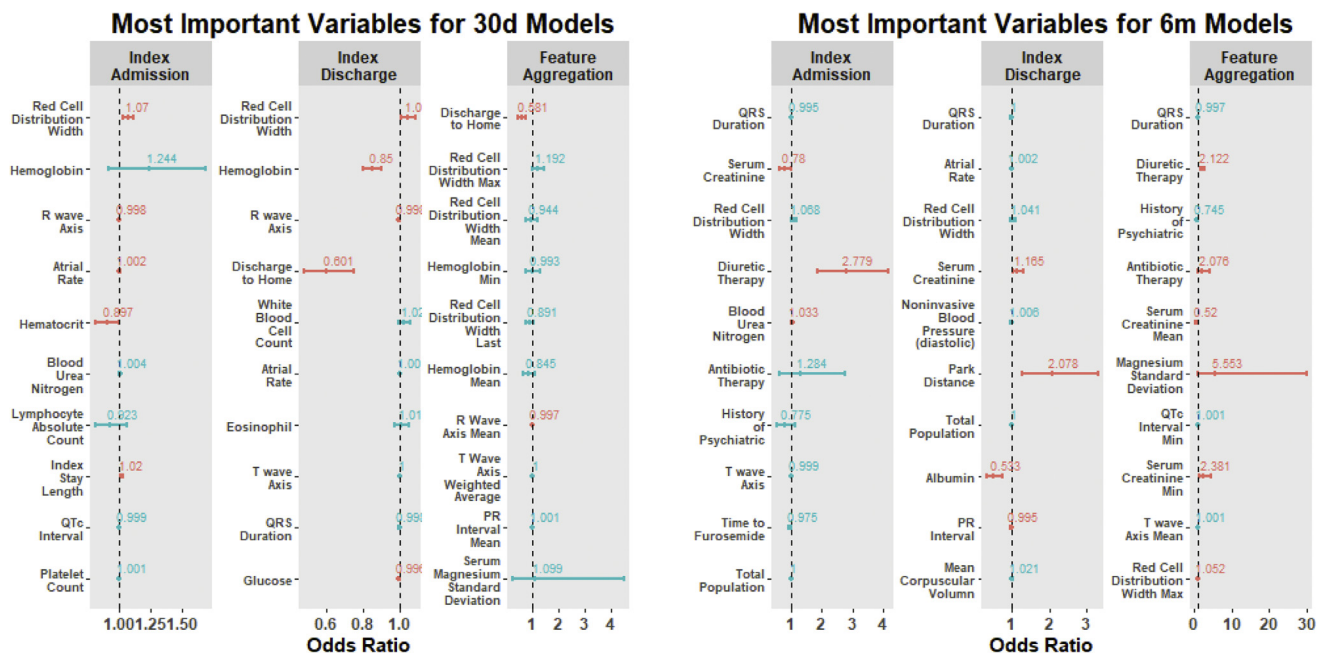| 30-Day outcomes			6-Month outcomes		
Index admission	Index discharge	Feature aggregated	Index admission	Index discharge	Feature aggregated
Serum potassium (1)	Total protein (1)	LA dimension weighted average (3)	Noninvasive blood pressure (1)	Serum chloride (1)	QTc interval mean (2)

Numbers in parentheses indicate the category in which each variable belongs, as follows: 1: hospital variable; 2: electrocardiography; 3: echocardiography; 4: social determinants of health; 5: discharge medication; 6: admission medication.

BNP = brain natriuretic peptide; LA = left atrium; PTT = partial thromboplastin time; WBC = white blood cell.

hospital readmissions. To our knowledge, this study incorporated more variables, both static and dynamic, in a feature-aggregated model than any other study to date. In addition, our analysis showed that certain variables (eg, hemoglobin level, red cell distribution width, ECG features) are consistent with previous studies showing high predictive value in HF risk assessment, reinforcing that these relationships may warrant further investigation.

Several predictive models have been developed in an attempt to estimate readmission risk among patients hospitalized with HF (Supplemental Table A.2). These models incorporate administrative or clinical variables using regression modeling, with limited discriminative ability. For example, Felker et al[13] evaluated 949 HF patients for 60-day mortality or the composite of death or rehospitalization at 60 days. They used a Cox proportional hazards model to identify variables that independently predicted risk and subsequently used a logistic regression model to predict the risk of death or rehospitalization. The c-statistic for the composite outcome was 0.68 (after bootstrapping). Keenan et al[14] used 283,919 hospitalizations in a derivation cohort to build a predictive model for the occurrence of 30-day readmissions. They

used a generalized linear model with a logit link function to construct the model, which had AUC = 0.60 for the primary outcome. In contrast, Yamokoski et al[15] enrolled 373 hospitalized patients with advanced HF and used logistic regression to predict the occurrence of rehospitalization or death within 6 months. The c-statistic for rehospitalization prediction was 0.519. These examples suggest that, despite a multitude of known risk factors, the actual prediction of HF rehospitalization is difficult and raise the possibility that important predictors of HF readmission are not represented in such models. Our investigation attempted to overcome these limitations by using a big data approach integrating static data (clinical and socioeconomic variables) with a multitude of imaging results (echocardiography, computed tomography, invasive coronary angiography), hemodynamic data, ECG, medications, and built environment (ie, human-made aspects of the environment including buildings, urban infrastructure, and parks for both physical and social human activities) data while utilizing state-of-the-art ML algorithms that have been shown to provide optimal predictive modeling effectiveness. Consequently, our index admission and index discharge models predicted

Figure 4   Quantification of the strength of association between the top variables for each model and 30-day (30d) or 6-month (6m) outcomes. Odds ratios (*red* indicates statistically significant; *blue* indicates nonsignificant) and associated 95% confidence intervals are shown.

30-day readmission and all-cause mortality with improved performance compared to the currently used HOSPITAL score for avoidable readmissions, while producing a high AUC compared to other investigations related to HF readmission prediction.[17,28,29] Of note, the HOSPITAL score predicted potentially avoidable readmission with a *c*-statistic of 0.72 (95% confidence interval 0.72–0.72) in a validation study by Donzé et al[27] using an international cohort of patients consecutively discharged from the medicine floor. The lower performance of the HOSPITAL score in this study may be due to inclusion of primarily HF patients, resulting in some elements of the HOSPITAL score, such as discharge from oncology service, being less relevant. Nonetheless, the HOSPITAL score is a well-validated score for hospital readmissions that is commonly used in clinical practice.

ML has been widely used across various medical specialties with the goal of producing highly accurate predictive modeling for specific events. ML is a field that draws on statistics, mathematics, and computer science to propose novel strategies for the construction of data-driven models from large datasets.[29,30] ML-based modeling predicts an outcome based on complicated and nonlinear relationships between certain variables and a specific outcome of interest, and can produce improved prediction models compared to those from systems that rely on expert-selected features. The present analysis showed that discharge to home (vs to rehabilitation or long-term care facility) ranked above both clinical and imaging variables, having the greatest predictive value for mortality and hospitalizations in the 30-day feature-aggregated model. This information is not routinely incorporated in HF risk prediction models and emphasizes that discharge planning could be an area for interventions, including increases in remote monitoring or outreach programs. Other important variables were chemistry based (eg, red cell distribution width, hemoglobin, hematocrit, blood urea nitrogen), which is consistent with a recent systematic review of 117 HF prediction models and a study reported by Angraal et al[31] using ML techniques to predict mortality and hospitalizations in patients with HF and preserved ejection fraction.[32] Relevant hemodynamic parameters were limited to echocardiographic variables, such as the E/E' ratio (a measure of left ventricular filling pressure), aortic valve area, and left atrial area, which could be a barometric measure of the chronicity and severity of mitral valvular disease, left ventricular systolic or diastolic dysfunction, or all aforementioned conditions. Likewise, QRS duration was highly predictive of unplanned readmission or all-cause mortality in almost every model, again possibly representing the extent of left ventricular dysfunction.[33] This is consistent with a recent study by Raghunath et al[34] showing that a deep neural network can predict 1-year all-cause mortality from ECG voltage–time traces (AUC 0.88), thus highlighting the need for further investigations to determine the relationship between ECG features and readmission or all-cause mortality. Invasive hemodynamic measures were not top predictors; this could be explained by the heterogeneity in indication for right heart catheterization, which may be

performed to establish the diagnosis of HF in individuals with ambiguous or conflicting signs and symptoms, whereas in others it could be performed to monitor and tailor therapy. Overall, the use of ML for predictive modeling within health care will potentially improve risk assessment and help streamline resources or guide new interventions. This study emphasizes the need to target patients discharged to their home.

Study limitations

In line with recent data, our findings suggest that short-term outcomes prediction in HF is a complex task that requires an innovative and comprehensive approach. Nevertheless, several limitations associated with the investigation should be noted. First, single-center data were collected retrospectively using EHR resources. However, inclusion criteria were defined requiring that patients have follow-up data in the EHR within 30 days of discharge in order to demonstrate that they received ongoing care at the institution. Future studies may consider utilizing the New York City Clinical Data Research Network, which includes multisite clinical records on patients across 20 New York City medical centers. However, such a database does not have the granular and longitudinal data of the CLEVER-HEART cohort and introduces interinstitutional variance in documentation and transformation practices. Second, the present analysis relied on hospital billing data to capture HF hospitalizations, which has implicit limitations in evaluating the risk of disease progression and acuity. Lastly, ML methods do not provide the ability to infer the direction of association between variables and a particular outcome. Our analysis partially averted this limitation by selecting the top predictors in the model and using them to construct a multivariate logistic regression model, which in turn provided a sense of both the strength and direction of the association between the particular variable and the primary outcome. However, use of a logistic regression model to elucidate the association between the highest-ranked variables may simplify complex relationships such as hemoglobin, which was shown in the Seattle Heart Failure Model to have a U-shaped relationship with mortality.[35] Nevertheless, despite these limitations, to the best of our knowledge the present study is the most comprehensive analysis, as characterized by its incorporation of numerous variables (beyond the traditional clinical and socioeconomic variables typically included in risk scores) as well as the utilization of primary HF patients and tailoring the outcome to unplanned readmissions or mortality.

Conclusion

The present investigation developed an improved and integrated model for the prediction of 30-day unplanned readmission or all-cause mortality in a contemporary cohort of primary HF patients. Several pathophysiological markers of importance in this model previously have been shown to be significant in other risk prediction models and reinforce the possible value of further investigation. The incorporation

of comprehensive EHR data and the utilization of novel ML algorithms could lead to the development of accurate predictive models that are disease-specific and tailored to a particular outcome of interest, leading to more efficient and cost-effective health care resource utilization.

Funding Sources

Disclosures

Dr Shaw is on the scientific advisory board for Covanos, Inc. Dr Al'Aref is supported by NIH 2R01 HL127661-05 and receives royalty fees from Elsevier. All other authors have reported that they have no conflicts relevant to the contents of this paper to disclose.

Appendix
Supplementary data

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.cvdhj.2020.07.004.

References

1. Benjamin EJ, Virani SS, Callaway CW, et al. Heart disease and stroke statistics—2018 update: a report from the American Heart Association. Circulation 2018;137:e67–e492.
2. Blecker S, Paul M, Taksler G, Ogedegbe G, Katz S. Heart failure–associated hospitalizations in the United States. J Am Coll Cardiol 2013;61:1259–1267.
3. Bergethon KE, Ju C, DeVore AD, et al. Trends in 30-day readmission rates for patients hospitalized with heart failure: findings from the Get With The Guidelines-Heart Failure Registry. Circ Heart Fail 2016;9:e002594.
4. Dharmarajan K, Hsieh AF, Lin Z, et al. Diagnoses and timing of 30-day readmissions after hospitalization for heart failure, acute myocardial infarction, or pneumonia. JAMA 2013;309:355–363.
5. Arora S, Lahewala S, Hassan Virk HU, et al. Etiologies, trends, and predictors of 30-day readmissions in patients with diastolic heart failure. Am J Cardiol 2017;120:616–624.
6. Chamberlain AM, Dunlay SM, Gerber Y, et al. Burden and timing of hospitalizations in heart failure: a community study. Mayo Clin Proc 2017;92:184–192.
7. O'Connor CM, Miller AB, Blair JE, et al. Causes of death and rehospitalization in patients hospitalized with worsening heart failure and reduced left ventricular ejection fraction: results from Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study with Tolvaptan (EVEREST) program. Am Heart J 2010;159:841–849.
8. Dunlay SM, Redfield MM, Weston SA, et al. Hospitalizations after heart failure diagnosis a community perspective. J Am Coll Cardiol 2009;54:1695–1702.
9. Suter LG, Li SX, Grady JN, et al. National patterns of risk-standardized mortality and readmission after hospitalization for acute myocardial infarction, heart failure, and pneumonia: update on publicly reported outcomes measures based on the 2013 release. J Gen Intern Med 2014;29:1333–1340.
10. Bradley EH, Curry L, Horwitz LI, et al. Hospital strategies associated with 30-day readmission rates for patients with heart failure. Circ Cardiovasc Qual Outcomes 2013;6:444–450.
11. Philbin EF, DiSalvo TG. Prediction of hospital readmission for heart failure: development of a simple risk score based on administrative data. J Am Coll Cardiol 1999;33:1560–1566.
12. Krumholz HM, Chen Y-T, Wang Y, Vaccarino V, Radford MJ, Horwitz RI. Predictors of readmission among elderly survivors of admission with heart failure. Am Heart J 2000;139:72–77.
13. Felker GM, Leimberger JD, Califf RM, et al. Risk stratification after hospitalization for decompensated heart failure. J Card Fail 2004;10:460–466.
14. Keenan PS, Normand SL, Lin Z, et al. An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. Circ Cardiovasc Qual Outcomes 2008;1:29–37.
15. Yamokoski LM, Hasselblad V, Moser DK, et al. Prediction of rehospitalization and death in severe heart failure by physicians and nurses of the ESCAPE trial. J Card Fail 2007;13:8–13.
16. Hammill BG, Curtis LH, Fonarow GC, et al. Incremental value of clinical data beyond claims data in predicting 30-day outcomes after heart failure hospitalization. Circ Cardiovasc Qual Outcomes 2011;4:60–67.
17. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. JAMA 2011;306:1688–1698.
18. Damen JA, Hooft L, Schuit E, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. BMJ 2016;353:i2416.
19. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. JAMA Cardiol 2017;2:204–209.
20. Golas SB, Shibahara T, Agboola S, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. BMC Med Inform Decis Mak 2018;18:44.
21. Sholle ET, Kabariti J, Johnson SB, et al. Secondary Use of Patients' Electronic Records (SUPER): an approach for meeting specific data needs of clinical and translational researchers. AMIA Annu Symp Proc 2018;2017:1581–1588.
22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–381.
23. Horwitz LI, Grady JN, Cohen DB, et al. Development and validation of an algorithm to identify planned readmissions from claims data. J Hosp Med 2015;10:670–677.
24. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). New York: Association for Computing Machinery; 2016. p. 785–794.
25. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operation characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845.
26. Greiner M. Two-graph receiver operating characteristic (TG-ROC): update version supports optimisation of cut-off values that minimise overall misclassification costs. J Immunol Methods 1996;191:93–94.
27. Donzé J, Aujesky D, Williams D, Schnipper JL. Potentially avoidable 30-day hospital readmissions in medical patients: derivation and validation of a prediction model. JAMA Intern Med 2013;173:632–638.
28. Ross JS, Mulvey GK, Stauffer B, et al. Statistical models and patient predictors of readmission for heart failure: a systematic review. Arch Intern Med 2008;168:1371–1386.
29. Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of machine learning techniques for heart failure readmissions. Circ Cardiovasc Qual Outcomes 2016;9:629–640.
30. Al'Aref SJ, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. Eur Heart J 2019;24:1975–1986.
31. Angraal S, Mortazavi BJ, Gupta A, et al. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. JACC Heart Fail 2020;1:12–21.
32. Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. JACC Heart Fail 2014;5:429–436.
33. Murkofsky RL, Dangas G, Diamond JA, Mehta D, Schaffer A, Ambrose JA. A prolonged QRS duration on surface electrocardiogram is a specific indicator of left ventricular dysfunction. J Am Coll Cardiol 1998;2:476–482.
34. Raghunath S, Ulloa Cerna AE, Jing L, et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. Nat Med 2020;26:886–891.
35. Levy W, Mozaffarian D, Linker DT, et al. The Seattle heart failure model. Circulation 2006;113:1424–1433.