Analysis of Machine Learning Techniques for Heart Failure Readmissions

Bobak J. Mortazavi, PhD; Nicholas S. Downing, MD; Emily M. Bucholz, MD, PhD; Kumar Dharmarajan, MD, MBA; Ajay Manhapra, MD; Shu-Xia Li, PhD; Sahand N. Negahban, PhD*; Harlan M. Krumholz, MD, SM*

Background—The current ability to predict readmissions in patients with heart failure is modest at best. It is unclear whether machine learning techniques that address higher dimensional, nonlinear relationships among variables would enhance prediction. We sought to compare the effectiveness of several machine learning algorithms for predicting readmissions.

Methods and Results—Using data from the Telemonitoring to Improve Heart Failure Outcomes trial, we compared the effectiveness of random forests, boosting, random forests combined hierarchically with support vector machines or logistic regression (LR), and Poisson regression against traditional LR to predict 30- and 180-day all-cause readmissions and readmissions because of heart failure. We randomly selected 50% of patients for a derivation set, and a validation set comprised the remaining patients, validated using 100 bootstrapped iterations. We compared C statistics for discrimination and distributions of observed outcomes in risk deciles for predictive range. In 30-day all-cause readmission prediction, the best performing machine learning model, random forests, provided a 17.8% improvement over LR (mean C statistics, 0.628 and 0.533, respectively). For readmissions because of heart failure, boosting improved the C statistic by 24.9% over LR (mean C statistic 0.678 and 0.543, respectively). For 30-day all-cause readmission, the observed readmission rates in the lowest and highest deciles of predicted risk with random forests (7.8% and 26.2%, respectively) showed a much wider separation than LR (14.2% and 16.4%, respectively).

Conclusions—Machine learning methods improved the prediction of readmissions after hospitalization for heart failure compared with LR and provided the greatest predictive range in observed readmission rates. (Circ Cardiovasc Qual Outcomes. 2016;9:900-0. DOI: 10.1161/CIRCOUTCOMES.116.003039.)

Key Words: computers ▶ heart failure ▶ machine learning ▶ meta-analysis ▶ patient readmission

High rates of readmission after hospitalization for heart failure impose tremendous burden on patients and the healthcare system.1–3 In this context, predictive models facilitate identification of patients at high risk for hospital readmissions and potentially enable direct specific interventions toward those who might benefit most by identifying key risk factors. However, current predictive models using administrative and clinical data discriminate poorly on readmissions.4–6 The inclusion of a richer set of predictor variables encompassing patients’ clinical, social, and demographic domains, while improving discrimination in some internally validated studies,7 does not necessarily markedly improve discrimination,8 particularly in the data set to be considered in this work. This richer set of predictors might not contain the predictive domain of variables required, but does represent a large set of data not routinely collected in other studies.

Another possibility for improving models, rather than simply adding a richer set of predictors, is that prediction might improve with methods that better address the higher order interactions between the factors of risk. Many patients may have risk that can only be predicted by modeling complex relationships between independent variables. For example, no...
WHAT IS KNOWN

• Prediction models for readmissions in heart failure are often modest, at best, at discriminating between readmitted and nonreadmitted patients.

WHAT THE STUDY ADDS

• This study compares popular machine learning methods for comparison against traditional techniques.
• This study introduces models that can select variables for us from a larger, comprehensive data set.
• This study shows the improvements in using machine learning methods and a comprehensively collected data set, to help readers understand factors that contribute to readmission and limitations to current methods in readmission prediction.

available variable may be adequately explanatory; however, interactions between variables may provide the most useful information for prediction.

Modern machine learning (ML) approaches can account for nonlinear and higher dimensional relationships between a multitude of variables that could potentially lead to an improved explanatory model.12,13 Many methods have emerged from the ML community that can construct predictive models using many variables and their rich nonlinear interactions.14–15 Three widely used ML approaches may provide utility for readmission prediction: random forest (RF), boosting, and support vector machines (SVM). The primary advantage of ML methods is that they handle nonlinear interactions between the available data with similar computational load. RF involves the creation of multiple decision trees that sort and identify important variables for prediction.14,17 Boosting algorithms harness the power of weaker predictors by creating combined and weighted predictive variables.18,19 This technique has been applied to preliminary learning work with electronic health records.20 SVM is a methodology that creates clearer separation of classes of variables using nonlinear decision boundaries, or hyperplanes, from complex, multidimensional data in possibly infinite dimensional spaces.21–26

In this study, we tested whether these ML approaches could predict readmissions for heart failure more effectively than traditional approaches using logistic regression (LR).7,9,27,28 We tested these strategies using data that included detailed clinical and sociodemographic information collected during the Tele-HF trial (Telemonitoring to Improve Heart Failure Outcomes),29 a National Institutes of Health–sponsored randomized clinical trial to examine the effect of automated telemonitoring on readmission after hospitalization for heart failure.10,29,30 This set included data extracted from medical record abstractions, hospital laboratory results, physical examination information, as well as quality of life, socioeconomic, and demographic information from initial patient surveys. The polarity of qualitative and categorical questions was altered if necessary to ensure that the lowest values reflect a strongly negative answer or missing data, and the highest values correspond to strongly positive answers (Data Supplement). In addition, we created dummy variables for each of the 236 features to indicate whether the value was missing or not (0 and 1 values).

Definition of Outcomes

We developed methods to predict 4 separate outcomes: (1) 30-day all-cause readmission; (2) 180-day all-cause readmission; (3) 30-day readmission because of heart failure; and (4) 180-day readmission because of heart failure. We trained predictive models for each of these outcomes and compared them with each other.

Predictive Techniques

We built models using both traditional statistical methods and ML methods to predict readmission, and compared model discrimination and predictive range of the various techniques. For traditional statistical methods, we used an LR model, and a Poisson regression. Three ML methods—RF, boosting, and SVM—were used for readmission prediction.

Logistic Regression

A recent study from Tele-HF used random survival forest methods to select from a comprehensive set of variables for Cox regression prediction models for readmissions in heart failure.16 That sort and identify important variables for prediction.14,17

Methods

Data Source

Data for this study were drawn from Tele-HF, which enrolled 1653 patients within 30 days of their discharge after an index hospitalization for heart failure.29 In addition to the clinical data from the index admission, Tele-HF used validated instruments to collect data on patients’ socioeconomic, psychosocial, and health status. This study collected a wide array of instrumented data, from comprehensive, qualitative surveys to detailed hospital examinations, including many pieces of data not routinely collected in practice, providing a larger exploratory set of variables that might provide information gain.

The primary outcome was all-cause readmission or mortality within 180 days of enrollment.29 A committee of physicians adjudicated each potential readmission to ensure that the event qualified as a readmission rather than another clinical encounter (eg, emergency department visit) and to determine the primary cause of the readmission. The comprehensive nature in which outcomes were tracked and determined across the various sites makes this a well-curated data set that can potentially leverage this information where other trials may not, as readmissions often occur at other institutions external to the study network.30 Results of the Tele-HF study revealed that outcomes were not significantly different between the telemonitoring and control arms in the primary analysis.29

Analytic Sample Selection

For this study, we included all patients whose baseline interviews were completed within 30 days of hospital discharge to ensure that the information was reflective of the time of admission. Of the 1653 enrolled patients, we excluded 36 who were readmitted or died before the interview, 574 whose interviews were completed after 30 days from discharge, and 49 who were missing data on ≥15 of the 236 baseline features to create a study sample of 1004 patients for the 30-day readmission analysis set. To create the 180-day readmission analysis set, we further excluded 27 patients who died before the end of the study and had no readmission events, leaving 977 patients in the sample.

Feature Selection

We used 472 variables (called features in the ML community) for input. We first gathered the full 236 baseline patient characteristics available in Tele-HF.29,30 This set included data extracted from medical record abstractions, hospital laboratory results, physical examination information, as well as quality of life, socioeconomic, and demographic information from initial patient surveys. The polarity of qualitative and categorical questions was altered if necessary to ensure that the lowest values reflect a strongly negative answer or missing data, and the highest values correspond to strongly positive answers (Data Supplement). In addition, we created dummy variables for each of the 236 features to indicate whether the value was missing or not (0 and 1 values).
analytic. The article had a comprehensive evaluation of the most predictive variables in the Tele-HF data set, using various techniques and various validation strategies. Using the variables selected in the article would provide the most accurate representation of an LR model on the Tele-HF data set for comparison purposes. Therefore, we used the same selected variables for our current study to compare model performance, as the current analysis is concerned with finding improved analytic algorithms for predicting 30-day readmissions rather than primarily with variable selection.

To verify that this LR model, which leverages the comprehensive variable selection technique in previous work, was the best model for comparison, we compared its performance against other LR models built using varied feature selection techniques. The first model selected variables for LR by lasso regularization, the next a forward, stepwise feature selection technique based on each variable’s likelihood ratio. All models were validated using 100 bootstrapped iterations; the former could not find a set of predictive variables, the latter varied in features chosen but selected, on average, 5 variables. The model built using features selected from the work of Krumholz et al outperformed the other techniques, when comparing mean C statistics. For a further detailed discussion of the other LR models built, we refer readers to the Data Supplement.

Comparison Techniques

Given the flexibility of nonlinear methods, the complexity of the desired models might overwhelm the available data, resulting in overfitting. Although all the available variables can be used in ML techniques such as RF and boosting, which are robust to this overfitting, we may require some form of feature selection to help prevent overfitting in less robust techniques like SVM. Further explanation of the predictive techniques, as well as the evaluation method, is provided in the Data Supplement and additionally in the textbook by Hastie et al.

Hierarchical Methods

To overcome the potential for overfitting in LR and SVM, we developed hierarchical methods with RF. Previous hierarchical methods used RF as a feature selection method because it best suited to a data set of high dimensionality with varied data types, to identify a subset of features to feed into methods such as LR and SVM. RF is well known to use out-of-bag estimates and an internal bootstrap to help reduce and select only predictive variables and avoid overfitting, similar to AdaBoost. RF is well known to be able to take varied data types and high-dimensional data and reduce to a usable subset, which we also verify through our bootstrapped cross-validation (through the use of a derivation and validation set). However, rather than to use the list of variables supplied by RF as the inputs to SVM or LR, the Tele-HF data set allowed us to use the probability predicted by RF as the inputs to SVM and to LR to create 2 new hierarchical methods.

The Tele-HF data set has comprehensive outcome information for all patients. We leveraged this by designing 2 prediction models using all of the aforementioned methods in a hierarchical manner. We used the RF algorithm as a base for this technique. The RF model, trained on all of the available features, produced a probability for many readmission events (eg, 0–12 events in this data set). These probabilities were then given to LR and SVM as inputs from which to build models. A detailed discussion of the method by which RF and SVM as well as RF and LR are combined together is in the Data Supplement.

Analytic Approach

For each of the predictive techniques above, we iterated the analyses illustrated in Figure 1 100×. To construct the derivation and validation data sets, we split the cohort into 2 equally sized groups, ensuring equal percentages of readmitted patients in each group. To account for a significant difference in numbers of patients who were readmitted and not readmitted in each group, we weighted the ML algorithms. The weight selected for the readmitted patients was the ratio of not-readmitted patients to readmitted patients in the derivation set.

The testing and selecting of appropriate weights are further detailed in the Data Supplement.

Once the derivation and validation sets were created, a traditional LR model was trained. We used SAS and the 5 most important variables as identified previously: blood urea nitrogen, glomerular filtration rate, sex, waist:hip ratio, and history of ischemic cardiomyopathy. The remaining techniques were created using the full raw data of 472 inputs for each patient and were trained on the different readmission outcome labels.

We trained models to predict either the 30-day readmission or 180-readmission outcome. Although we supplied the same input data to each method, we varied the outcome information provided. Because the 180-day readmission set contains more outcomes information, including the total number of readmission events during the trial, it is possible that it might be easier to predict 180-day readmission, given the propensity of some patients to be readmitted multiple times. To provide a range of modeling for the final binary prediction (readmitted/not readmitted), we ran 5 distinct training methods based on the labels provided to the algorithm. For 30-day readmission prediction, we first generated 3 different training models with the same baseline data, but 3 different outcomes: 30-day binary outcomes, 180-day binary outcomes, and 180-day counts of readmissions. We then used the predictive models created by these 3 training methods to predict 30-day readmission outcomes in the validation cohort. Similarly, to test 180-day readmission, we generated 2 different training methods with same baseline data but different outcomes, namely, 180-day binary outcomes and 180-day counts of readmissions. A detailed description of how these methods vary from each other is available in the Data Supplement.

We ran the models generated on the validation set and calculated the area under the receiver operating characteristics curve (C statistic), which provided a measure of model discrimination. The analysis was run 100× in order to provide robustness over a potentially poor random split of patients and to generate a mean C statistic with a 95% confidence interval (CI). We also evaluated the risk-stratification abilities of each method. The probabilities of readmission generated over the 100 iterations were then sorted into deciles. Finally, we calculated the observed readmission rate for each decile to determine the predictive value of the algorithms.

These models should be used prospectively, to provide clinicians with decision-making points. For each iteration, we calculated the positive predictive value (PPV), sensitivity, specificity, and f-score, a common measurement in ML, which is calculated as

$$f\text{-score} = 2 * \frac{\text{PPV} * \text{Sensitivity}}{\text{PPV} + \text{Sensitivity}}$$

The f-score provides a balance between PPV and sensitivity, similar to a C statistic, but at designated thresholds on the ROC curve. The data-driven decision threshold for prospective modeling measurements was that which maximized the f-score.

Furthermore, to test whether a narrowed focus of prediction could improve discrimination, we conducted the same analyses using heart failure–only readmission instead of all-cause readmission.

All ML analyses were developed in R. The list of R packages used in this study is included in the Data Supplement. The Human Investigation Committee at the Yale School of Medicine approved the study protocol.

Results

Baseline characteristics of patients in 30- and 180-day analytic samples are detailed in Table 1. In both analytic samples, the proportion of women and blacks was ≈40%. Ninety percent of patients had New York Heart Association class II or III heart failure, with ≈70% having left ventricular ejection fraction of <40%. A history of hypertension (77%) and diabetes mellitus (45%) was common. The 30-day all-cause readmission rate was 17.1%, whereas the 180-day all-cause readmission rate was 48.9%. No sex-specific or race-specific differences were found.
Thirty-Day All-Cause Model Discrimination
The discrimination of the different predictive models for 30-day all-cause readmission represented by C statistic values is illustrated in Figure 2A. LR had a low C statistic of 0.533 (95% CI, 0.527–0.538). Boosting on the input data had the highest C statistic (0.601; 95% CI, 0.594–0.607) in 30-day binary outcome with 30-day training case. Boosting also had the highest C statistic for the 30-day binary outcome with 180-day binary training (0.613; 95% CI, 0.607–0.618). For the 30-day outcomes with 180-day counts training, the RF technique had the highest C statistic (0.628; 95% CI, 0.624–0.633).

One Hundred Eighty–Day All-Cause Model Discrimination
The discrimination values for different models in predicting the 180-day readmission are illustrated in Figure 2B. LR again showed a low C statistic (0.574; 95% CI, 0.571–0.578) for the 180-day binary case. The RF into SVM hierarchical method had the highest achieved C statistic across all methods (0.654; 95% CI, 0.650–0.657) in 180-day binary outcome and 180-day count case (0.649; 95% CI, 0.648–0.651).

Thirty-Day All-Cause Model Predictive Range
Deciles of the probabilities/responses generated by the algorithms are plotted against the observed rate of readmission in each decile in Figure 3. For each training version of the method, we plotted the values from the method and training version with the highest C statistic (ie, best RF from the three 30-day training scenarios). In Figure 2A, the RF has the same mean C statistic and similar CI as the RF into LR hierarchical model. In Figure 3, for RF, the readmission rates were markedly higher at the tenth decile of probabilities (26.2%) than the first decile of probabilities (14.2%) showing a higher predictive range than LR (14.2% and 16.4%, respectively), whereas for the RF into LR hierarchical model, the actual readmission rates varied only from 10.6% to 19.0% from the first to the tenth decile, despite the same mean C statistic for RF and RF into LR.

One Hundred Eighty–Day All-Cause Model Predictive Range
When the predicted probabilities of each algorithm were separated into deciles and plotted against the real readmission rate per decile, RF into SVM had the largest difference in readmission rate from the first decile (31.0%) to the tenth decile.
(69.4%), whereas LR had a readmission rate of 40.5% in the first decile and 60.1% in the tenth decile (Figure 4).

**Readmission Because Of Heart Failure Discrimination**

As illustrated in Figure 5A and 5B, for readmissions because of heart failure, the LR model again had a low C statistic for the 30-day binary case (0.543; 95% CI, 0.536–0.550) and the 180-day binary case (0.566; 95% CI, 0.562–0.570). Boosting had the best C statistic for the 30-day binary-only case (0.615; 95% CI, 0.607–0.622) and for the 30-day with 180-day binary case training (0.678; 95% CI, 0.670–0.687). The highest C statistic for other prediction cases were RF for the 30-day with 180-day counts case training (0.669; 95% CI, 0.661–0.676); RF into SVM for the 180-day counts case (0.651; 95% CI, 0.646–0.656).

**Readmission Because Of Heart Failure Predictive Range**

When the deciles of risk prediction were plotted against the observed readmission rate, RF and boosting each had the biggest differences between the first and tenth deciles of risk (1.8–11.9% and 1.4–12.2%, respectively). Although the hierarchical methods of RF into LR and RF into SVM had similar C statistics (0.654; 95% CI, 0.645–0.663 and 0.656; 95% CI, 0.648–0.664, respectively) in Figure 6, only RF into LR clearly identifies low- and high-risk groups. The risk stratification for the 180-day case because of heart failure follows closely with C statistics, similar to that for the 180-day all-cause case (Figure 7).

### Table 1. Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>30 d</th>
<th>180 d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All, n (%)</strong></td>
<td>Readmission, n (%)</td>
<td>All, n (%)</td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>1004 (100.0)</td>
<td>149 (14.8)</td>
</tr>
<tr>
<td><strong>Median age, y (SD)</strong></td>
<td>62 (15.7)</td>
<td>62 (15.9)</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td>415 (41.3)</td>
<td>58 (38.9)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>White</strong></td>
<td>507 (50.5)</td>
<td>84 (56.4)</td>
</tr>
<tr>
<td><strong>Black</strong></td>
<td>393 (39.1)</td>
<td>55 (36.9)</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>104 (10.4)</td>
<td>10 (6.7)</td>
</tr>
<tr>
<td><strong>New York Heart Association</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Class I</strong></td>
<td>56 (5.6)</td>
<td>7 (4.7)</td>
</tr>
<tr>
<td><strong>Class II</strong></td>
<td>515 (51.3)</td>
<td>85 (57.0)</td>
</tr>
<tr>
<td><strong>Class III</strong></td>
<td>355 (35.4)</td>
<td>52 (35.0)</td>
</tr>
<tr>
<td><strong>Class IV</strong></td>
<td>58 (5.8)</td>
<td>2 (1.3)</td>
</tr>
<tr>
<td><strong>Missing</strong></td>
<td>20 (1.9)</td>
<td>3 (2.0)</td>
</tr>
<tr>
<td><strong>Medical history</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LVEF % &lt;40</strong></td>
<td>687 (68.4)</td>
<td>99 (66.4)</td>
</tr>
<tr>
<td><strong>Hypertension</strong></td>
<td>771 (76.8)</td>
<td>116 (77.9)</td>
</tr>
<tr>
<td><strong>Diabetes mellitus</strong></td>
<td>450 (44.8)</td>
<td>57 (38.3)</td>
</tr>
<tr>
<td><strong>Myocardial infarction</strong></td>
<td>257 (25.6)</td>
<td>47 (31.5)</td>
</tr>
<tr>
<td><strong>Stroke</strong></td>
<td>96 (9.6)</td>
<td>15 (10.1)</td>
</tr>
<tr>
<td><strong>Ischemic cardiomyopathy</strong></td>
<td>235 (23.4)</td>
<td>44 (29.5)</td>
</tr>
<tr>
<td><strong>Clinical values, mean/SD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Albumin</strong></td>
<td>3.32 (0.58)</td>
<td>3.25 (0.61)</td>
</tr>
<tr>
<td><strong>Blood urea nitrogen</strong></td>
<td>25.2 (17.8)</td>
<td>28.7 (20.3)</td>
</tr>
<tr>
<td><strong>Creatinine</strong></td>
<td>1.40 (0.77)</td>
<td>1.53 (0.80)</td>
</tr>
<tr>
<td><strong>Hemoglobin</strong></td>
<td>12.3 (1.94)</td>
<td>12.0 (1.80)</td>
</tr>
<tr>
<td><strong>Glomerular filtration rate</strong></td>
<td>58.5 (27.4)</td>
<td>52.8 (24.6)</td>
</tr>
<tr>
<td><strong>Potassium</strong></td>
<td>4.08 (0.57)</td>
<td>4.19 (0.58)</td>
</tr>
</tbody>
</table>

LVEF indicates left ventricular ejection fraction.

*All values are mean (SD) unless noted.*
Improvements Over LR

Table 2 shows the percentage improvement in prediction of the best-trained model for each ML technique over LR for each outcome predicted. RF achieved a 17.8% improvement in discrimination over LR, the best 30-day all-cause readmission improvement. Similarly, the hierarchical method of RF into SVM achieved a 13.9% improvement for 180-day all-cause readmission; boosting achieved a 24.9% improvement for 30-day readmission because of heart failure; and the hierarchical method of RF into SVM achieved a 16.1% improvement over LR for 180-day readmission because of heart failure.

Prediction Results

Table 3 shows the prospective prediction results of the best model for 30-day all-cause readmission, 180-day all-cause readmission, 30-day readmission because of heart failure, and 180-day readmission because of heart failure. Thirty-day all-cause readmission had a PPV of 0.22 (95% CI, 0.21–0.23), sensitivity of 0.61 (0.59–0.64), and a specificity of 0.61 (0.58–0.63) at a maximal f-score of 0.32 (0.31–0.32); 180-day all-cause readmission had a PPV of 0.51 (0.51–0.52), a sensitivity of 0.92 (0.91–0.93), and a specificity of 0.18 (0.16–0.21) at a maximal f-score of 0.66 (0.65–0.66); 30-day readmission because of heart failure had a PPV of 0.15 (0.13–0.16), a sensitivity of 0.45 (0.41–0.48), and a specificity of 0.79 (0.76–0.82) at a maximal f-score of 0.20 (0.19–0.21); 180-day readmission because of heart failure had a PPV of 0.51 (0.50–0.51), a sensitivity of 0.94 (0.93–0.95), and a specificity of 0.15 (0.13–0.17) at a maximal f-score of 0.66 (0.65–0.66).

The 30-day predictions, in general, were better at identifying negative cases, whereas the 180-day predictions were better able to correctly identify positive cases.

Discussion

In our study to model risk of readmissions in a sample of patients hospitalized with heart failure, the ML analytic techniques we applied improved discrimination modestly compared with the traditional method of LR. These ML models also provided better identification of groups at low and high risk for readmission by increasing predictive range compared with LR. These results are consistent with our hypothesis that ML methods can leverage complex higher-level interactions among a multitude of variables to improve discrimination and
predictive range with respect to heart failure outcomes compared with traditional linear models. We used readmissions among patients with heart failure as a test case to examine the advantage of ML algorithms over traditional statistical analysis to see whether such an approach could be applied to other outcomes prediction and predictive range problems.

Models developed from ML algorithms to improve discrimination and predictive range can be evaluated using

Figure 3. Deciles of algorithm risk versus readmission rates (%) for the best 30-d all-cause models for each method. The y axis presents the observed readmission rates in each decile, the x axis the ordered deciles of risk predicted. SVM indicates support vector machine.

Figure 4. Deciles of algorithm risk versus readmission rates (%) for the best 180-d all-cause models for each method. The y axis presents the observed readmission rates in each decile, the x axis the ordered deciles of risk predicted. SVM indicates support vector machine.
several measures or methods. The modest improvement in
discrimination represented by C statistics across the vari-
ous ML methods, while promising, is ultimately insufficient.
Furthermore, the models selected, at a data-driven threshold
that optimizes the f-score, optimized the correct classification
of negative cases but had a large number of false positives,
as seen in the low PPV and f-score. This threshold decision
would be varied if further cost information was inputted to
the model or the selected threshold optimized on another fac-
tor, such as reducing false alarms. We found that although
discriminatory power was similar across certain methods,
the range in observed events among deciles of predicted risk
varied greatly. Analyzing multiple aspects of models pro-
vides a stronger understanding of those models. With regard
to 30-day all-cause readmission case, RF better differentiates
low- and high-risk groups compared with all other methods
including the hierarchical RF into LR model. Despite similar
discrimination (mean C statistic) between the 2 methods, RF
actually has a better predictive range. In other words, RF bet-
ter discriminates patients in low- and high-risk strata, whereas
hierarchical RF into LR better discriminates patients in the
middle deciles of risk. This suggests that a combination of
various approaches, with possible deployment of different
methods at different strata of risk, can improve discrimination
of the overall model and may improve the predictive range.

There have been some suggestions in the literature about
how to improve predictive range of risk strata and overall
discrimination. Wiens et al.26 showed that use of time-based
information can improve their predictive range and overall
discrimination when predicting the risk of *Clostridium dif-
icile* infection in hospitalized patients. In an earlier study
examining heart failure mortality prediction using cytokine
measurements, use of the baseline cytokine measurements
alone did not significantly improve discrimination.34 However,
the addition of serial cytokine measurements over follow-up
time to the predictive model improved discrimination for
1-year mortality.34 Thus, perhaps leveraging serially collected
telemonitoring data on daily health and symptoms of patients
could improve discrimination and predictive range of our
predictions.

The general experience with LR is that readmissions after
index hospitalizations for heart failure are more difficult to pre-
dict than mortality outcomes. In a meta-analysis, Ouwerkerk
et al12 analyzed 117 heart failure outcome prediction models

![Figure 5. Forest plots for C statistics and 95% confidence intervals (CIs) for each method with respect to prediction of 30- and 180-d heart failure (HF) readmission (A and B). Diamond represents C statistic, and the line represents the 95% CI. Color coding for model training: red: trained in the 30-d binary case only; green: trained with 30- and 180-d binary outcome; blue: trained with the 180-d counts case. SVM indicates support vector machine.](http://circoutcomes.ahajournals.org/Downloaded from)
with 249 variables, the average C statistic to predict mortality was 0.71 and hospitalization was 0.63. Tele-HF study data were collected to identify and track heart failure–related symptoms, and we find that narrowing our prediction to heart failure–only readmissions improves the discrimination and predictive range of each method.

Ultimately, despite the increase in a wide array of data, from quality of life questionnaires to hospital examinations,
Limitations
The discrimination and range of prediction are likely affected by several limitations of the Tele-HF data set. First, the number of patients is small in comparison to the number of variables measured per patient. As a result, while rich in terms of patient features, the small number of patients makes the data set vulnerable to overfitting. Furthermore, many variables included in models with stronger C statistics from previous studies are not present in Tele-HF, which could affect performance. Time-to-event analysis conducted on the Tele-HF data set produced a C statistic of 0.65. This finding suggests there are measurable domains of variables not analyzed in this study that might be more predictive in all patients with heart failure. However, with the moderate-at-best C statistics, it is possible that the best set of features in predicting risk in patients with heart failure has not yet been collected.

The interpretation of the results can be varied depending on the particular data set and purpose of modeling. For example, the prediction threshold selected at the maximal f-score may not be desirable for a particular intervention designed when considering a model, such as when used in a prospective fashion with the false alarm rate in mind. Furthermore, the numeric ranges of the probabilities generated by the algorithms do not necessarily match, and thus the thresholds considered across each method vary. Although the output of the algorithm provides a predicted probability of readmission, the value in this is not the number provided but where this lies in the predictive range of the method and, thus, whether it can be classified as low or high risk, where the observed rates are considered. This is a result of each method basing its distance (and thus probability) of observations in the model on different metrics, resulting in a different calculation of a final probability. If these probabilities would be desirable for direct use, the values could be scaled to match comparable intervals across methods.

The trial arm of Tele-HF may provide further insights into the value of time-to-event analysis by permitting inclusion of the daily telemonitoring responses by patients in the intervention arm as a time-series signal to be processed for time-dependent risk models. To determine the accuracy of additional methods, a direct comparison in the predictions of each individual patient (both risk and outcome) needs to be made between methods rather than net reclassification index work, which has not been shown to be effective in many cases at understanding and improving model discrimination performance. Finally, to address the viability of the model and the clinical predictors it is built on, an external validation cohort should be used. However, as the clinical predictors collected in studies are often varied, it becomes important to distinguish unique predictors in internal validation, as well as understand the impact each predictor has on the models for better interpretation. As this is a post hoc analysis of a randomized controlled trial, the generalizability of such models is a limitation. Nevertheless, the goal of this study was to compare methods, not to introduce a new model. Therefore, there is little reason to think that a population-based sample would yield a different result. Although the data from such a trial span a diverse set of hospitals and patients, and is well curated, once the understanding of such strengths and weaknesses of ML methods is understood, they should be extended.

| Table 2. Percent Improvement of Machine Learning Methods Over Baseline, Traditional Logistic Regression |
|-----------------------------------------------|------------------|------------------|------------------|------------------|
| 30-d all-cause readmission prediction          | −3.75            | 17.8             | 15.0             | 17.8             |
| 180-d all-cause readmission prediction         | 7.14             | 13.8             | 11.0             | 13.8             |
| 30-d heart failure–only readmission prediction| −5.71            | 23.2             | 24.9             | 20.4             |
| 180-d heart failure–only readmission prediction| 2.83             | 16.1             | 15.7             | 16.1             |

| Table 3. Prospective Prediction Results (Mean and 95% Confidence Intervals) |
|-----------------------------------------------|------------------|------------------|------------------|------------------|
| Positive Predictive Value                    | Sensitivity      | Specificity      | F-Score          |
| 30-d all-cause readmission                   | 0.22 (0.21–0.23) | 0.61 (0.59–0.64) | 0.61 (0.58–0.63) | 0.32 (0.31–0.32) |
| 180-d all-cause readmission                  | 0.51 (0.51–0.52) | 0.92 (0.91–0.93) | 0.18 (0.16–0.21) | 0.66 (0.65–0.66) |
| 30-d heart failure–only readmission          | 0.15 (0.13–0.16) | 0.45 (0.41–0.48) | 0.79 (0.76–0.82) | 0.20 (0.19–0.21) |
| 180-d heart failure–only readmission         | 0.51 (0.50–0.51) | 0.94 (0.93–0.95) | 0.15 (0.13–0.17) | 0.66 (0.65–0.66) |
to data derived from other sources as well, including electronic health records, which will likely affect results because of a consideration of varied data availability and a more heterogeneous patient population.

Conclusions
The results of our study support the hypothesis that ML methods, with the ability to leverage all available data and their complex relationships, can improve both discrimination and range of prediction over traditional statistical techniques. The performance of the ML methods varies in complex ways, including discrimination and predictive range. Future work needs to focus on further improvement of predictive ability through advanced methods and more discerning data, so as to facilitate better targeting of interventions to subgroups of patients at highest risk for adverse outcomes.

Sources of Funding
Dr Dharmarajan was supported by grant K23AG048331 from the National Institute on Aging, Bethesda, MD, and the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program, New York, NY. Dr Krumholz was supported by grant U01 HL105270 (Center for Cardiovascular Outcomes Research at Yale University) from the NHLBI in Bethesda, MD. Dr Manhapra was supported by the Research in Addiction Medicine Scholars (RAMS) Program, R25DA033211 from the National Institute on Aging, Bethesda, MD, and the American Heart Association, Bethesda, MD. Dr Manhapra was supported by grant K23AG048331 from the National Institute on Aging, Bethesda, MD. Dr Dharmarajan was supported by grant K23AG048331 from the National Institute on Aging, Bethesda, MD. Dr Krumholz was chair of a cardiac scientific advisory board for UnitedHealth, is the recipient of research agreements for Clover Health, Inc. Dr. Krumholz is chair of a cardiac scientific advisory board for the Centers for Medicare & Medicaid Services to develop and maintain performance measures. Dr. Dharmarajan serves on a scientific advisory board for Clover Health, Inc. Dr. Krumholz is chair of a cardiac scientific advisory board for UnitedHealth, is the recipient of research agreements from Medtronic and Johnson and Johnson (Jansen), to develop methods of clinical trial data sharing, and is the founder of Hugo, a personal health information platform. The other authors report no conflicts.

Disclosures
Drs Dharmarajan and Krumholz work under contract with the Centers for Medicare & Medicaid Services to develop and maintain performance measures. Dr Dharmarajan serves on a scientific advisory board for Clover Health, Inc. Dr Krumholz is chair of a cardiac scientific advisory board for UnitedHealth, is the recipient of research agreements from Medtronic and Johnson and Johnson (Jansen), to develop methods of clinical trial data sharing, and is the founder of Hugo, a personal health information platform. The other authors report no conflicts.

References


Analysis of Machine Learning Techniques for Heart Failure Readmissions
Bobak J. Mortazavi, Nicholas S. Downing, Emily M. Bucholz, Kumar Dharmarajan, Ajay Manhapra, Shu-Xia Li, Sahand N. Negahban and Harlan M. Krumholz

Circ Cardiovasc Qual Outcomes. published online November 8, 2016;
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2016 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/early/2016/11/08/CIRCOUTCOMES.116.003039
Free via Open Access

Data Supplement (unedited) at:
http://circoutcomes.ahajournals.org/content/suppl/2016/11/09/CIRCOUTCOMES.116.003039.DC1
http://circoutcomes.ahajournals.org/content/suppl/2016/11/09/CIRCOUTCOMES.116.003039.DC2

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org/subscriptions/
SUPPLEMENTAL MATERIAL

Mortazavi: Analysis of Machine Learning Techniques for Heart Failure Readmissions

Supplemental Methods

Data Set Creation

Variables

This section details the creation of the features, the alignment of particular features, and the missing features. In particular, Table S1 shows the list of all baseline variables collected in the Tele-HF trial. Removing Patient ID, the 236 remaining variables each had a duplicate binary variable created to indicate whether that value was missing or not (e.g., AGE_MISSING would be a binary variable with “yes” if the age was missing for that given patient and “no” otherwise).

Data Alignment

The values in the variables are of three types, either binary, continuous, or categorical. However, due to the combination of multiple questionnaire types, the values assigned to the binary and categorical answers do not match. For example, one categorical value may have a numeric range from 1 to 5 where 1 is the worst answer and 5 is the best. The very next question may have five options as well but order them 0 to 4 where 4 is the worst. As a result, before any algorithms were run, all of the variables were ordered to increase as the answer improved (aligning the intensities as much as possible). The worst answer was given a value of 1 and the best answer would go up from there, 5 in most cases but up to 7 or 8 in many. Thus, the missing values would result in 0 being assigned. In this case the missing value provides no weight to the machine learning algorithm and can be considered either its own category or aligned with a truly negative response. This was chosen to integrate well with future endeavors on the telemonitoring data where the hypothesis to be tested is that missing data correlates strongly with negative responses and adverse outcomes.
Predictive Methods

Setting up the Outcome

Before we discuss the algorithms in further detail we should further outline how we arrived at the conclusion that we should weight our readmission cases heavily (equal to the proportion of not-readmitted vs. readmitted cases) versus the many other possibilities run in the 30-day prediction case (since the 180-day case had a roughly equal number of readmits to non-readmits). We ran the following iterations of dealing with class imbalance. We first applied all the algorithms with no weighting and found that the probability of being predicted as a readmit was quite low. In order to balance the sets, we tried a number of techniques. We downsampled the non-readmit cases to be equal to the number of readmit cases in the training set, but found that the number of training samples was simply too small to create an effective model. We then upsampl e the readmitted cases to be equal to the number of non-readmit cases. This improved some of the algorithms and not all. The final approach we took was to vary our weighting of the readmit cases versus the non-readmit cases. Comparing the results of no weighting, downsampling, and upsampling, we observed that the weights setting the two sets to roughly equal seemed to be the best (which it was). To further confirm this observation, we varied the weight across a range from no weighting to twice as strong as the proportion of non-readmits to readmits. The rest of the methods described were run for each of these cases but the final outputs presented in the paper concern only the weighted examples.

Logistic Regression (LR) Comparisons

Three different LR models were built in order to validate the strength of the model presented in the main text, as well as to validate its use as a comparative technique to the machine learning (ML) models built. The first such model, based on GLMNET\textsuperscript{1}, was originally used for LR with Lasso regularization for feature selection, but the cross-validated LR in this model deemed no variables worthy of including in the final model so it produced a C-statistic of 0.5 since the only feature in the model was the intercept.
The second such model used the 236 Tele-HF variables and computed a forward, stepwise selection based upon the likelihood ratio of each variable. Forward selection was chosen because in certain high-dimensional datasets, the results are actually considered closer to the optimal solution than a Lasso regularization as in GLMNET. The method was run until no variable’s likelihood had a p-value < 0.01. This technique produced a model, in 100 bootstrapped iterations, with a mean C-statistic of 0.524 with a 95% confidence interval over the 100 iterations of (0.518-0.529). Further, the model selected on average only 4.79 (95% CI: 4.5-5.1) variables. These variables selected were different in each iteration, with 80 of the 236 variables being selected across the 100 iterations. The variables selected are listed in Table S2. Note that the frequently-selected variables do not match those from Krumholz et al., indicating their valuation of variable importance results in the selection of more appropriate predictors.

The third such model used the full 472 variables created for this paper, including the dummy missing variables. Features were forward-selected similarly, and the method produced a C-statistic of 0.518 (0.512-0.524) with on average 5.53 (5.2-5.9) variables. Incidentally, this method selected 85 of the 472 variables, none of which were the missing dummy variables. These variables served only to affect the likelihood calculations and selections of the variables considered in the second method above, and increasing the number of variables beyond 5 actually lowered the C-statistic. As a result, the method chosen in the main text, based upon a prior study that comprehensively reviewed and evaluated the predictive capabilities of each variable in the Tele-HF dataset, produces the best and fairest comparison technique for the work.

*Inputs and Outputs to Each Method*

This section will cover in greater detail the machine learning algorithms considered and how they were coded in R for replication. SAS was used (proc logistic) to model the logistic regression as explained in the main text. The remainder of the techniques were coded in R.

The first ML technique, Poisson Regression, uses the input data as well as the total number of readmissions to create a predictive model based upon the propensity to be readmitted.
This technique is classically used to predict a range of counts (e.g., the number of readmission events) it can also be used for comparing the binary case of readmitted/not readmitted, has a built in feature ordering and selection technique, and outputs the variables associated with the predictive model along with a propensity value for prediction. This value is given to the pROC package along with the ground truth labels to determine the ROC curve and area under the curve estimate.³

The second technique, Random Forests⁴ (RF), uses a series of decision trees on the input data to predict a final outcome by considering the result of each decision tree. The decision trees can be trained to output a binary decision or a prediction on the range of readmissions. Further, the ordering of features in the trees gives a selection of the most important features used for prediction. Finally, RF can be trained using the number of readmissions or the binary label of readmitted/not readmitted and can output probabilities of a binary prediction or a multiclass prediction (e.g., probability of each particular number of readmissions). For each of the test cases we had RF output the probabilities of being within each class. For the binary readmission case, the probability of being in class 1 (readmitted) was used for pROC calculations. For the counts of readmissions, we are given a matrix of probabilities where each column indicates each count, namely, 0 for no readmissions, 1 for 1 readmission, 2 for 2 readmissions and so on. We tested a combination of factors to add probabilities for a final 0 or 1 prediction. We tried the final 0 prediction probability being only the column associated with class 0. We then added the probability of being in class 1 with class 0 and called this a probability of not being readmitted, and so forth through all of the probabilities. When being fed into LR or support vector machines (SVM), it was these matrix of probabilities that were supplied as inputs. The results presented in the paper indicate the best form of this approach, where class 0 was considered no readmissions and the probabilities of all the other classes were added together to form the probability of being readmitted.
The third technique, that also contains a form of feature ranking, is Boosting. Boosting attempts to take a series of weak classifiers (e.g., tree classifiers based upon a single feature) that only classify pieces of the data well, and re-weigh them to develop an overall strong classifier. This iterative technique then builds a strong classifier as a weighted linear combination of the results of these weak classifiers, to output a binary outcome. We used the ADA package\(^5\) to test this method. This package has the flexibility of providing two loss functions, exponential and logistic, as well as the different kind of boosting techniques, discrete (also known as AdaBoost), real (for RealBoost), and gentle (for GentleBoost). The results presented in the paper were a summary of the best form of Boosting calculated by the algorithms.

The final ML technique considered is the SVM, implemented in R by the package e1071, of which the SVM implementation is known as LibSVM\(^6\). An SVM is a supervised learning model that leverages higher dimensional spaces to attempt to determine separation between the classes being trained. Unlike the previous methods, however, a feature selection algorithm must be run prior to the SVM to select the most relevant features. Such selection algorithms can be take many forms, and in this work, will be provided by RF, where the output from the RF models (the matrix of probabilities) will serve as the inputs to the SVM. Both linear and radial basis function kernels (RBF) were used but in each case the RBF algorithm outperformed the linear kernel. In some cases, the linear kernel was unable to converge on a solution in the given number of iterations.

Again, in all cases, instead of looking at a final response output, we looked at the probabilities of being readmitted generated by the algorithm, then supplied that to the pROC package to vary thresholds of readmitted/not readmitted and generate an ROC curve for us.

**Using RF with Other Methods**

RF is a method that is well-suited to a dataset of high dimensionality with a number of mixed types.\(^7\) RFs build decision trees where each node is split by a single chosen variable. This variable is selected by using out-of-bag estimates and bootstrapping to measure error, correlation
to other variables, and strength of prediction, to pick the best variables as well as ensure the model does not overfit. For this reason, RFs are often used in situations with a high-dimensional, varied dataset, to serve as a feature selection technique for other models as well as its own model. For this reason, this method is often used to select features, with the top importance features used to train methods such as SVM.

This work, similarly, leverages the ability of the RF method to evaluate a large set of variables and develop a predictive model without overfitting. However, rather than taking the selected variables, which might be of different types (e.g. continuous and categorical), RF can produce the probability of multiple events. For example, rather than have a binary yes/no prediction of readmission, RF can create a regression for the number of readmissions (e.g. 0-12) and provide a probability of each readmission count. These probabilities (13 of them in this case), are then provided as inputs to SVM and LR, methods that are at risk of overfitting if all the variables were to be provided to them. Thus, the hierarchical models created avoid overfitting by leveraging RFs, which avoid overfitting by using an internal bootstrapping and out-of-bag errors to ensure this. Further, the 100 bootstrapped iterations and the accuracy produced help verify that this is the case experimentally.

Creating Deciles of Risk

Similar to the ROC creation, each iteration of responses was also split into deciles using the R quantile function. Each iteration the boundary values from the predicted responses are taken for the deciles and a mean boundary value plus 95% confidence intervals around those boundaries are calculated. The total set of responses are also then combined and split into deciles to show the fraction of times correct across the cross-validated samples. We verified that the decile boundaries, created by the entire list of responses, falls within the 95% confidence interval calculated by each method for its particular range of responses. These were then used to give the observed readmission rates as detailed in the manuscript and results. Table S3 gives an example of this for 30-day all-cause readmissions and the boundaries presented by RF.
Cohort Results

For the additional patients eliminated from the analysis because they died before being readmitted, we evaluated characteristics versus the cohort used for training. The analytic sample did not differ significantly from those who were excluded for various reasons (Table S4). Further, as the study excluded a large number of the original 1653 participants, we have listed the differences between the included cohort (n=1004) and excluded cohort (n=649) (Table S5). While many of the values are similar it might be interesting to further analyze their differences in future work. Finally, for the included patients, the percent missing for each variable is plotted in Figure S1. The complete missing information can be found in Table S6. In particular, the high rates of missing values for a large number of variables seems to be in large part as a result of incomplete questionnaires leading to summary scores that could not be calculated. While no individual patient within the cohort selected has a large rate of missing data, it does seem that the variables missing are consistent across all of the patients, which improves the likelihood that imputation is not causing a large effect on the outcome.

SUPPLEMENTAL REFERENCES


SUPPLEMENTAL FIGURE LEGEND

Figure S1. Variable missing rates for all variables with a rate of missing greater than 3% across the 1004 patient cohort.
## SUPPLEMENTAL TABLES

### Table S1. The 236 Features Used in the Tele-HF Analysis

<table>
<thead>
<tr>
<th>Feature</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient Information</strong></td>
<td></td>
</tr>
<tr>
<td>Age Over 90</td>
<td>Age greater than 90?</td>
</tr>
<tr>
<td>Age</td>
<td>Age of Patient</td>
</tr>
<tr>
<td>Age De-Identified</td>
<td>De-identified Age</td>
</tr>
<tr>
<td>Payor Type 1</td>
<td>Insurance: Commercial/PPO</td>
</tr>
<tr>
<td>Payor Type 3</td>
<td>Insurance: HMO</td>
</tr>
<tr>
<td>Payor Type 2</td>
<td>Insurance: Medicaid</td>
</tr>
<tr>
<td>Payor Type 5</td>
<td>Insurance: Medicare</td>
</tr>
<tr>
<td>Payor Type 6</td>
<td>Insurance: None/Self-Pay</td>
</tr>
<tr>
<td>Payor Type 7</td>
<td>Insurance: Other</td>
</tr>
<tr>
<td>Payor Type 8</td>
<td>Insurance: Unknown</td>
</tr>
<tr>
<td>Payor Type 4</td>
<td>Insurance: VA</td>
</tr>
<tr>
<td>Telemonitoring</td>
<td>Is patient assigned to the telemonitoring group?</td>
</tr>
<tr>
<td>HF HOSP</td>
<td>Is the number of times admitted due to heart failure an estimate?</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Is the patient Hispanic, non-Hispanic, or unknown?</td>
</tr>
<tr>
<td>Rx1_0</td>
<td>Medication: On ACE Inhibitor or ARB</td>
</tr>
<tr>
<td>Rx2_0</td>
<td>Medication: On ARA</td>
</tr>
<tr>
<td>Rx3_0</td>
<td>Medication: On Beta Blocker</td>
</tr>
<tr>
<td>Rx4_0</td>
<td>Medication: On Digoxin</td>
</tr>
<tr>
<td>Rx5_0</td>
<td>Medication: On Loop Diuretics?</td>
</tr>
<tr>
<td>RACE_4</td>
<td>Race: American Indian/Alaska Native</td>
</tr>
<tr>
<td>RACE_3</td>
<td>Race: Asian</td>
</tr>
</tbody>
</table>
RACE_2  Race: Black/African-American
RACE_5  Race: Native Hawaiian/Pacific Islander
RACE_6  Race: Other
RACE_7  Race: Unknown
RACE_1  Race: White/Caucasian
HOSP   Number of hospital visits in past year
Admitted Number of times admitted to hospital in past year
HF_HOSP Number of times admitted to hospital in past year for heart failure
PATIENT_SCALE Patient has their own weight scale
PATID  Patient ID in Trial
SEX    Patient Sex

Hospitalization Information
Symptom: Chest Cause of Hospitalization: Chest Pain
Symptom: Fatigue Cause of Hospitalization: Fatigue
Symptom: Heart Cause of Hospitalization: Irregular Heart Beat or Palpitations
Symptom: Other Cause of Hospitalization: Other
Symptom: Breath Cause of Hospitalization: Shortness of breath
Symptom: Swelling Cause of Hospitalization: Swelling

Patient History
Connective Tissue History: Connective Tissue Disease
Coronary Artery History: Coronary Artery Disease
Diabetes History: Diabetes
Diabetes: Organ History: Diabetes with end organ damage
AIDS History: Does patient have AIDS?
Hemiplegia History: does patient have hemiplegia?
<table>
<thead>
<tr>
<th>Condition</th>
<th>History:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor</td>
<td>Does patient have history of tumor?</td>
</tr>
<tr>
<td>Apnea</td>
<td>Does patient have sleep apnea?</td>
</tr>
<tr>
<td>Liver Disease Rating</td>
<td>If patient has liver disease, is it mild or moderate/severe?</td>
</tr>
<tr>
<td>Dialysis</td>
<td>On Dialysis</td>
</tr>
<tr>
<td>Leukemia</td>
<td>Patient has a history of leukemia</td>
</tr>
<tr>
<td>Liver Disease</td>
<td>Patient has a history of liver disease</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Patient has a history of lymphoma</td>
</tr>
<tr>
<td>Metastatic Tumor</td>
<td>Patient has a history of metastatic tumors</td>
</tr>
<tr>
<td>Pacemaker</td>
<td>Patient has a permanent pacemaker</td>
</tr>
<tr>
<td>Chronic renal failure</td>
<td>Patient has chronic renal failure</td>
</tr>
<tr>
<td>Cardio Resync</td>
<td>Patient has had cardio resynchronization therapy</td>
</tr>
<tr>
<td>Cerebrovascular Disease</td>
<td>Patient has had cerebrovascular disease</td>
</tr>
<tr>
<td>Chronic Pulmonary</td>
<td>Patient has had chronic pulmonary disease</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>Patient has history of hypercholesterolemia</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Patient has history of hypertension</td>
</tr>
<tr>
<td>Drugs</td>
<td>Patient has history of illicit drug use</td>
</tr>
<tr>
<td>Ischemic Cardiomyopathy</td>
<td>Patient has history of ischemic cardiomyopathy</td>
</tr>
<tr>
<td>Peptic Ulcer</td>
<td>Patient has peptic ulcer disease</td>
</tr>
<tr>
<td>PVD</td>
<td>Patient has peripheral vascular disease</td>
</tr>
<tr>
<td>Aortic</td>
<td>Patient has severe aortic or mitral valve disease</td>
</tr>
<tr>
<td>Disease Management</td>
<td>Patient is in another disease management program</td>
</tr>
<tr>
<td>Oxygen</td>
<td>Patient uses Oxygen at home</td>
</tr>
<tr>
<td>Chronic Renal: Type</td>
<td>Patient with chronic renal failure has either mild or moderate/severe</td>
</tr>
<tr>
<td>Dementia</td>
<td>Patient has dementia</td>
</tr>
</tbody>
</table>
### AICD
History: Prior AICD implantation

### MI
History: Prior MI

### TIA
History: Prior TIA

#### Laboratory Values and Physical Exams

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin</td>
<td>Laboratory Value: Albumin</td>
</tr>
<tr>
<td>Blood Urea Nitrogen</td>
<td>Laboratory Value: Blood Urea Nitrogen</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>Laboratory Value: Body Mass Index</td>
</tr>
<tr>
<td>BNP</td>
<td>Laboratory Value: Brain Natriuretic Peptide</td>
</tr>
<tr>
<td>CREATININE</td>
<td>Laboratory Value: Creatinine</td>
</tr>
<tr>
<td>DIASTOLIC_BP</td>
<td>Laboratory Value: Diastolic BP</td>
</tr>
<tr>
<td>GFR</td>
<td>Laboratory Value: Glomular Filtration Rate</td>
</tr>
<tr>
<td>HEIGHT</td>
<td>Laboratory Value: Height</td>
</tr>
<tr>
<td>HEMOCRIT</td>
<td>Laboratory Value: Hemocrit</td>
</tr>
<tr>
<td>HEMOGLOBIN</td>
<td>Laboratory Value: Hemoglobin</td>
</tr>
<tr>
<td>HIP</td>
<td>Laboratory Value: Hip circumference</td>
</tr>
<tr>
<td>JVD_MEASURE</td>
<td>Method by which JVD was calculated</td>
</tr>
<tr>
<td>JVD</td>
<td>Laboratory Value: Jugular Venous Distension</td>
</tr>
<tr>
<td>LVEF_METHOD</td>
<td>Laboratory Value: method by which LVEF was calculated</td>
</tr>
<tr>
<td>NT_PRO_BNP</td>
<td>Laboratory Value: N-terminal pro b-type Natriuretic Peptide</td>
</tr>
<tr>
<td>NYHA</td>
<td>Laboratory Value: NYHA Class</td>
</tr>
</tbody>
</table>
| LOW_COG                    | Laboratory Value: Patient's Folstein score was less than or equal to 24
| LVEF_PERCENT               | Laboratory Value: Percentage of LVEF                   |
| PITTING_EDEMA              | Laboratory Value: Pitting Edema (yes, no, or unsure)    |
| POTASSIUM                  | Laboratory Value: Potassium                            |
S3_PRESENCE Laboratory Value: Presence of S3 (yes, no, unsure)
PULMONARY_RALES Laboratory Value: Pulmonary Rales (Base, Above, or Clear)
RESP_RATE Laboratory Value: Respiratory Rate
RESTING_HR Laboratory Value: Resting Heart Rate
SYSTOLIC_BP Laboratory Value: Systolic Blood Pressure
TEMP Laboratory Value: Temperature
WAIST Laboratory Value: Waist Circumference
LVEF40 Laboratory Value: Was patient's LVEF under 40%?
WEIGHT Laboratory Value: Weight
Folstein Folstein mini mental status exam score
Hemorrhagic Hemorrhagic type: None, yes, or no
Surveys
Contribute Did someone other than the patient complete the baseline survey?
Help Complete Did the patient have someone help complete surveys?
Work for Pay Does patient currently work for pay?
Alone Does patient live alone?
Doctor Does the patient currently have a primary care provider?
Insurance Does the patient have insurance?
Smoke Does the patient smoke? And if so how often?
Financial Financially, how comfortable is patient's household on patient's income?
Difficult How difficult is it for the patient to receive health care?
Follow Up How good is the patient's doctor at following up?
Smoke QT How many cigarettes in the past 30 days?
Doc Days How many days has patient been seeing current doctor?
<table>
<thead>
<tr>
<th>Question</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doc Mos</td>
<td>How many months has patient been seeing current doctor?</td>
</tr>
<tr>
<td>Dependent</td>
<td>How many people are dependent upon patient's income?</td>
</tr>
<tr>
<td>Doc Yrs</td>
<td>How many years has patient been seeing current doctor?</td>
</tr>
<tr>
<td>Medcosts</td>
<td>How often did patient avoid taking medication due to costs?</td>
</tr>
<tr>
<td>Smoke Age</td>
<td>How old was patient when she/he started smoking regularly?</td>
</tr>
<tr>
<td>Knowledge</td>
<td>How well does patient know own health?</td>
</tr>
<tr>
<td>Patient Scale: Source</td>
<td>If patient has a weight scale, did trial provide it?</td>
</tr>
<tr>
<td>Visits per week</td>
<td>If patient has been visited by home health care, how many visits per week?</td>
</tr>
<tr>
<td>Visits QT</td>
<td>If patient has been visited by home health care, how many visits?</td>
</tr>
<tr>
<td>Health</td>
<td>In general, how does the patient feel about her/his own health?</td>
</tr>
<tr>
<td>Visits</td>
<td>In past 3 months has patient been visited by home health care?</td>
</tr>
<tr>
<td>Avoided</td>
<td>In past year, has patient avoided getting health care because of cost?</td>
</tr>
<tr>
<td>Burden</td>
<td>In Past year, how much of a financial burden have medical costs been?</td>
</tr>
<tr>
<td>Studies</td>
<td>Is patient involved in other studies?</td>
</tr>
<tr>
<td>Religion</td>
<td>Is religion a source of strength for the patient?</td>
</tr>
<tr>
<td>INHOSPITAL</td>
<td>Was the patient's interview/screen conducted in a hospital?</td>
</tr>
<tr>
<td>EDUCATION</td>
<td>What is the highest level of education the patient has completed?</td>
</tr>
<tr>
<td>INCOME</td>
<td>What is the patient's total household income?</td>
</tr>
<tr>
<td>ENROLL_LANGUAGE</td>
<td>What language did the patient enroll in?</td>
</tr>
<tr>
<td>HOME</td>
<td>Where does the patient currently live?</td>
</tr>
<tr>
<td>ESSI</td>
<td></td>
</tr>
<tr>
<td>ESSI: Close</td>
<td>ESSI: How often does patient have contact with someone close to</td>
</tr>
</tbody>
</table>
them?

ESSI: How often does patient have someone she/he can count on to listen?

ESSI: How often does patient have someone to help with chores?

ESSI: How often does patient have someone who can provide emotional support?

ESSI: How often does patient have someone who is able to provide good advice?

ESSI: How often does patient have someone who shows them affection?

ESSI: Summary Score of the ESSI Survey

KCCQ

KCCQ_WORK: KCCQ: How much does heart failure affect working or chores?

KCCQ_VISITING: KCCQ: How much does your heart failure affect visiting family?

KCCQ_HOBBIES: KCCQ: How much has heart failure affected your hobbies?

KCCQ_INTIMATE: KCCQ: How much has heart failure affected your intimate relationships?

KCCQ_LIFE: KCCQ: How satisfied with the rest of life would patient be with this level of heart failure?

KCCQ_CALL: KCCQ: how sure is patient in knowing who to call if heart failure gets worse?

KCCQ_UNDERSTAND: KCCQ: How well do you understand how to keep from getting worse?

KCCQ_SYMPTOMS: KCCQ: In past 2 weeks, have your symptoms changed?

KCCQ_LIMITED: KCCQ: In past 2 weeks, how much has heart failure limited your
<table>
<thead>
<tr>
<th>Question ID</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>KCCQ_FATIGUEBOTHER</td>
<td>KCCQ: In past 2 weeks, how often has fatigue bothered you?</td>
</tr>
<tr>
<td>KCCQ_FATIGUE</td>
<td>KCCQ: In past 2 weeks, how often has fatigue limited you?</td>
</tr>
<tr>
<td>KCCQ_YARDWORK</td>
<td>KCCQ: In past 2 weeks, how often has heart failure limited yard work?</td>
</tr>
<tr>
<td>KCCQ_DISCOURAGED</td>
<td>KCCQ: In past 2 weeks, how often has patient felt discouraged?</td>
</tr>
<tr>
<td>KCCQ_SHORTNESSBOTHER</td>
<td>KCCQ: In past 2 weeks, how often has shortness of breath bothered you?</td>
</tr>
<tr>
<td>KCCQ_SHORTNESS</td>
<td>KCCQ: In past 2 weeks, how often has shortness of breath limited you?</td>
</tr>
<tr>
<td>KCCQ_SLEEPING</td>
<td>KCCQ: In past 2 weeks, how often have you slept sitting up?</td>
</tr>
<tr>
<td>KCCQ_SWELLING</td>
<td>KCCQ: In past 2 weeks, how often have you woken up with swelling?</td>
</tr>
<tr>
<td>KCCQ_BATHING</td>
<td>KCCQ: In past two weeks, how limited by heart failure has bathing been?</td>
</tr>
<tr>
<td>KCCQ_DRESSING</td>
<td>KCCQ: In past two weeks, how limited by heart failure has dressing been?</td>
</tr>
<tr>
<td>KCCQ_JOGGING</td>
<td>KCCQ: In past two weeks, how limited by heart failure has jogging been?</td>
</tr>
<tr>
<td>KCCQ_STAIRS</td>
<td>KCCQ: In past two weeks, how limited by heart failure has stair climbing been?</td>
</tr>
<tr>
<td>KCCQ_WALKING</td>
<td>KCCQ: In past two weeks, how limited by heart failure has walking one block been?</td>
</tr>
<tr>
<td>KCCQ_BOTHER</td>
<td>KCCQ: In past two weeks, how much has swelling bothered you?</td>
</tr>
<tr>
<td>KCCQ_PHYSICAL</td>
<td>KCCQ: Physical Limitations Summary Score</td>
</tr>
</tbody>
</table>
KCCQ_QUALITY  KCCQ: Quality of Life Summary Score
KCCQ_EFFICACY  KCCQ: Self-Efficacy Summary Score
KCCQ_SOCIAL  KCCQ: Social Limitations Summary Score
KCCQ_SUMMARY  KCCQ: Summary Score
KCCQ_BURDEN  KCCQ: Symptom Burden Summary Score
KCCQ_FREQUENCY  KCCQ: Symptom Frequency Summary Score
KCCQ_STABILITY  KCCQ: Symptom Stability Summary Score
KCCQ_SYMPSCORE  KCCQ: Symptom Summary Score

Morisky

MORISKY_FORGOT  Morisky: Has patient forgotten to take medicine?
MORISKY_STOP  Morisky: Has patient stopped taking medication when feeling better?
MORISKY_WORSE  Morisky: Has patient stopped taking medication when feeling worse?
MORISKY_CARELESS  Morisky: Is Patient careless about taking medicine?
MISSMORISKY  Patient's Morisky Miss Summary Score
MORISKY  Patient's Morisky Summary Score

PHQ9

PHQ9_INTEREST  PHQ9: In past 2 weeks, how often had litter interest or pleasure doing things?
PHQ9_TIRED  PHQ9: In past two weeks, how often feeling tired?
PHQ9_SLOW  PHQ9: In past two weeks, how often moving or speaking slower than usual?
PHQ9_SLEEPING  PHQ9: In past two weeks, how often trouble sleeping?
PHQ9_FAILURE  PHQ9: Over last two weeks, how often felt bad or a failure?
PHQ9_DEPRESSED: PHQ9: Over last two weeks, how often felt down or depressed?
PHQ9: Over last two weeks, how often felt would be better off dead?
PHQ9: Over last two weeks, how often troubled by poor appetite?
PHQ9: Over last two weeks, how often troubled concentrating?
PHQ9: Summary Score

PSS

PSS_DIFFICULTY: PSS: In past month, difficulties were piling up
PSS_YOURWAY: PSS: in past month, felt things were going patient's way
PSS_CONTROL: PSS: In past month, felt unable to control important things?
PSS: In past month, how confident can handle personal problems?
PSS: Summary Score

REALM

REALM_NONE: REALM: could not read any terms
REALM_ALLERGIC: REALM: could read Allergic?
REALM_ANEMIA: REALM: could read Anemia?
REALM_COLITIS: REALM: could read Colitis?
REALM CONSTIPATION: REALM: could read Constipation?
REALM_DIRECTED: REALM: could read Directed?
REALM_FAT: REALM: could read Fat?
REALM_FATIGUE: REALM: could read Fatigue?
REALM_FLU: REALM: could read Flu?
REALM_JAUNDICE: REALM: could read Jaundice?
REALM_OSTEOPOROSIS
REALM: could read Osteoporosis?

REALM_PILL
REALM: could read Pill?

REALM_SCORE
REALM: Summary Score

SCHFI
SCHFI: How sure are you that you can do something to make symptoms better?

SCHFI_BETTER
SCHFI: How sure are you that you can judge how serious symptoms are?

SCHFI_SERIOUS
SCHFI: How sure are you that you can judge what makes symptoms better?

SCHFI_JUDGE
SCHFI: How sure are you that you can notice changes in health?

SCHFI
SCHFI: Summary Score

WARE
WARE_NEED
WARE: Can get medical care when patient needs it?

WARE_FINANCIAL
WARE: Can get medical care without financial setback?

WARE_EXPLAIN
WARE: Doctor explains things well?

WARE_TIME
WARE: Doctor spends plenty of time with patient?

WARE_FRIENDLY
WARE: Doctor treats patient in friendly manner?

WARE_IGNORE
WARE: Doctors ignores what patient tells them?

WARE_OFFICE
WARE: Doctor's office has everything patient needs?

WARE_DOUBTS
WARE: Doubts about how your doctor is treating you?

WARE_SPECIALIST
WARE: Easy access to medical specialist when needed?

WARE_CAREFUL
WARE: How careful is your doctor to check everything?

WARE_DISSATISFIED
WARE: How dissatisfied with things in your medical care are you?
WARE_APPOINTMENT  WARE: how hard is it to get an appointment?
WARE_CARE  WARE: How perfect do you feel your medical care is?
WARE_1  WARE: intermediate score
WARE_2  WARE: intermediate score
WARE_3  WARE: intermediate score
WARE_4  WARE: intermediate score
WARE_5  WARE: intermediate score
WARE_6  WARE: intermediate score
WARE_7  WARE: intermediate score
WARE_EMERGENCY  WARE: Medical care takes too long for emergency treatment?
WARE_HURRY  WARE: Providers hurry too much when treating patient?
WARE  WARE: Summary Score
WARE_AFFORD  WARE: With care being received now, how well do you feel you can afford care?
WARE_DOCTOR  WARE: Your doctor acts too business like?
Table S2. Variables Selected in the Forward, Stepwise Selection Method for LR

<table>
<thead>
<tr>
<th>Feature</th>
<th>Number of Times Selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALBUMIN_</td>
<td>1</td>
</tr>
<tr>
<td>METASTATIC_TUMOR_</td>
<td>5</td>
</tr>
<tr>
<td>CORONARY_ARTERY_</td>
<td>3</td>
</tr>
<tr>
<td>KCCQ_BOTHER_</td>
<td>3</td>
</tr>
<tr>
<td>ALONE_</td>
<td>1</td>
</tr>
<tr>
<td>WARE_1_</td>
<td>1</td>
</tr>
<tr>
<td>PHQ9_INTEREST_</td>
<td>3</td>
</tr>
<tr>
<td>CHRONIC_RENAL_FAILURE_</td>
<td>1</td>
</tr>
<tr>
<td>DEPENDENT_</td>
<td>1</td>
</tr>
<tr>
<td>WARE_EXPLAIN_</td>
<td>1</td>
</tr>
<tr>
<td>BLOOD_UREA_NITROGEN_</td>
<td>6</td>
</tr>
<tr>
<td>KCCQ_EFFICACY_</td>
<td>6</td>
</tr>
<tr>
<td>KCCQ_DISCOURAGED_</td>
<td>1</td>
</tr>
<tr>
<td>PATIENT_SCALE_SOURCE_</td>
<td>1</td>
</tr>
<tr>
<td>WARE_</td>
<td>3</td>
</tr>
<tr>
<td>DIABETES_</td>
<td>1</td>
</tr>
<tr>
<td>DOCYRS_</td>
<td>2</td>
</tr>
<tr>
<td>RACE_6_</td>
<td>1</td>
</tr>
<tr>
<td>PHQ9_</td>
<td>1</td>
</tr>
<tr>
<td>CONNECTIVE_TISSUE_</td>
<td>2</td>
</tr>
<tr>
<td>KCCQ_QUALITY_</td>
<td>1</td>
</tr>
<tr>
<td>KCCQ_STAIRS_</td>
<td>4</td>
</tr>
<tr>
<td>KCCQ_UNDERSTAND_</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Count</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>ENROLL_LANGUAGE</td>
<td>2</td>
</tr>
<tr>
<td>KCCQ_FREQUENCY</td>
<td>3</td>
</tr>
<tr>
<td>WARE_EMERGENCY</td>
<td>1</td>
</tr>
<tr>
<td>CREATININE</td>
<td>2</td>
</tr>
<tr>
<td>JVD_MEASURE</td>
<td>1</td>
</tr>
<tr>
<td>BURDEN</td>
<td>17</td>
</tr>
<tr>
<td>RELIGION</td>
<td>9</td>
</tr>
<tr>
<td>MORISKY_FORGOT</td>
<td>4</td>
</tr>
<tr>
<td>KCCQ_JOGGING</td>
<td>12</td>
</tr>
<tr>
<td>KCCQ_BURDEN</td>
<td>5</td>
</tr>
<tr>
<td>ESSI_CHORES</td>
<td>6</td>
</tr>
<tr>
<td>REALM_DIRECTED</td>
<td>1</td>
</tr>
<tr>
<td>MEDCOSTS</td>
<td>2</td>
</tr>
<tr>
<td>KCCQ_VISITING</td>
<td>2</td>
</tr>
<tr>
<td>CHRONIC_PULMONARY</td>
<td>5</td>
</tr>
<tr>
<td>ADMITTED</td>
<td>28</td>
</tr>
<tr>
<td>SCHFI</td>
<td>3</td>
</tr>
<tr>
<td>REALM_OSTEOPOROSIS</td>
<td>1</td>
</tr>
<tr>
<td>KCCQ_SHORTNESSBOTHER</td>
<td>2</td>
</tr>
<tr>
<td>KCCQ_STABILITY</td>
<td>1</td>
</tr>
<tr>
<td>KCCQ_WORK</td>
<td>4</td>
</tr>
<tr>
<td>FINANCIAL</td>
<td>1</td>
</tr>
<tr>
<td>DISEASE_MANAGEMENT</td>
<td>1</td>
</tr>
<tr>
<td>WARE_SPECIALIST</td>
<td>2</td>
</tr>
<tr>
<td>PAYOR_8</td>
<td>2</td>
</tr>
<tr>
<td>Variable</td>
<td>Count</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>KCCQ_WALKING_</td>
<td>1</td>
</tr>
<tr>
<td>GFR_</td>
<td>2</td>
</tr>
<tr>
<td>KCCQ_SWELLING_</td>
<td>24</td>
</tr>
<tr>
<td>RACE_3_</td>
<td>6</td>
</tr>
<tr>
<td>Rx1_0_</td>
<td>1</td>
</tr>
<tr>
<td>INSURANCE_</td>
<td>2</td>
</tr>
<tr>
<td>PSS_DIFFICULTY_</td>
<td>1</td>
</tr>
<tr>
<td>PAYOR_6_</td>
<td>2</td>
</tr>
<tr>
<td>KCCQ_SUMMARY_</td>
<td>5</td>
</tr>
<tr>
<td>OXYGEN_</td>
<td>1</td>
</tr>
<tr>
<td>Rx2_0_</td>
<td>1</td>
</tr>
<tr>
<td>DOCMOS_</td>
<td>1</td>
</tr>
<tr>
<td>HEALTH_</td>
<td>2</td>
</tr>
<tr>
<td>EDUCATION_</td>
<td>1</td>
</tr>
<tr>
<td>PEPTIC_ULCER_</td>
<td>1</td>
</tr>
<tr>
<td>LIVER_DISEASE_RATING_</td>
<td>1</td>
</tr>
<tr>
<td>KCCQ_HOBBIES_</td>
<td>1</td>
</tr>
<tr>
<td>KCCQ_YARDWORK_</td>
<td>5</td>
</tr>
<tr>
<td>HF_HOSP_</td>
<td>13</td>
</tr>
<tr>
<td>KCCQ_PHYSICAL_</td>
<td>4</td>
</tr>
<tr>
<td>KCCQ_SYMPSCORE_</td>
<td>1</td>
</tr>
<tr>
<td>SCHFI_HEALTH_</td>
<td>1</td>
</tr>
<tr>
<td>TUMOR_</td>
<td>2</td>
</tr>
<tr>
<td>VISITEDQT_</td>
<td>2</td>
</tr>
<tr>
<td>KCCQ_FATIGUE_</td>
<td>2</td>
</tr>
<tr>
<td>Variable</td>
<td>Count</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>HELP_COMPLETE_</td>
<td>3</td>
</tr>
<tr>
<td>POTASSIUM_</td>
<td>1</td>
</tr>
<tr>
<td>MORISKY_WORSE_</td>
<td>1</td>
</tr>
<tr>
<td>MORISKY_STOP_</td>
<td>4</td>
</tr>
<tr>
<td>KCCQ_SYMPTOMS_</td>
<td>4</td>
</tr>
<tr>
<td>PRIOR_AICD_</td>
<td>7</td>
</tr>
<tr>
<td>HF_HOSP_EST_</td>
<td>24</td>
</tr>
</tbody>
</table>
Table S3. Deciles and Confidence Intervals for 30-day All-Cause Readmission Responses from RF

<table>
<thead>
<tr>
<th>Deciles</th>
<th>Overall Boundary</th>
<th>Mean (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.392</td>
<td>0.393 (0.390-0.396)</td>
</tr>
<tr>
<td>2</td>
<td>0.418</td>
<td>0.420 (0.417-0.423)</td>
</tr>
<tr>
<td>3</td>
<td>0.438</td>
<td>0.439 (0.436-0.442)</td>
</tr>
<tr>
<td>4</td>
<td>0.456</td>
<td>0.456 (0.453-0.459)</td>
</tr>
<tr>
<td>5</td>
<td>0.472</td>
<td>0.472 (0.469-0.475)</td>
</tr>
<tr>
<td>6</td>
<td>0.488</td>
<td>0.487 (0.484-0.490)</td>
</tr>
<tr>
<td>7</td>
<td>0.504</td>
<td>0.503 (0.500-0.506)</td>
</tr>
<tr>
<td>8</td>
<td>0.524</td>
<td>0.522 (0.519-0.525)</td>
</tr>
<tr>
<td>9</td>
<td>0.550</td>
<td>0.548 (0.545-0.551)</td>
</tr>
<tr>
<td>10</td>
<td>1.00</td>
<td>1.00 (1.00-1.00)</td>
</tr>
</tbody>
</table>
### Table S4. Excluded Patients with Death but No Readmission from 180-Day Analysis Set

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Excluded</th>
<th>Included</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>N</td>
<td>27 (100.0)</td>
<td>977 (100.0)</td>
</tr>
<tr>
<td>Median age (SD)</td>
<td>66.5 (12.2)</td>
<td>62 (15.7)</td>
</tr>
<tr>
<td>Females</td>
<td>12 (44.4)</td>
<td>403 (41.2)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>16 (59.3)</td>
<td>491 (50.3)</td>
</tr>
<tr>
<td>African American</td>
<td>8 (29.6)</td>
<td>385 (39.4)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (11.1)</td>
<td>101 (10.3)</td>
</tr>
<tr>
<td>New York Heart Association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class I</td>
<td>2 (7.41)</td>
<td>54 (5.5)</td>
</tr>
<tr>
<td>Class II</td>
<td>15 (55.6)</td>
<td>500 (51.2)</td>
</tr>
<tr>
<td>Class III</td>
<td>8 (29.6)</td>
<td>347 (35.5)</td>
</tr>
<tr>
<td>Class IV</td>
<td>1 (3.70)</td>
<td>57 (5.8)</td>
</tr>
<tr>
<td>Missing</td>
<td>1 (3.70)</td>
<td>19 (2.0)</td>
</tr>
<tr>
<td>Medical History</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LVEF† % &lt; 40</td>
<td>19 (70.4)</td>
<td>668 (68.4)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>19 (70.4)</td>
<td>752 (77.0)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>11 (40.7)</td>
<td>439 (44.9)</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>7 (25.9)</td>
<td>250 (25.6)</td>
</tr>
<tr>
<td>Stroke</td>
<td>4 (14.8)</td>
<td>92 (9.4)</td>
</tr>
<tr>
<td>Ischemic</td>
<td>7 (25.9)</td>
<td>228 (23.3)</td>
</tr>
</tbody>
</table>
Cardiomyopathy

**Clinical Values**

*(Mean/SD)*

<table>
<thead>
<tr>
<th>Test</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin</td>
<td>3.5 (0.37)</td>
<td>3.31 (0.53)</td>
</tr>
<tr>
<td>Blood Urea Nitrogen</td>
<td>36.4 (26.5)</td>
<td>26.3 (16.8)</td>
</tr>
<tr>
<td>Creatinine</td>
<td>1.70 (1.21)</td>
<td>1.45 (0.72)</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>11.4 (1.75)</td>
<td>12.4 (1.94)</td>
</tr>
<tr>
<td>Glomerular Filtration Rate</td>
<td>49.3 (22.7)</td>
<td>58.8 (27.4)</td>
</tr>
<tr>
<td>Potassium</td>
<td>3.93 (0.59)</td>
<td>4.08 (0.57)</td>
</tr>
</tbody>
</table>

*All values in tables are mean (standard deviation) unless noted.

†LVEF: Left Ventricular Ejection Fraction.
### Table S5. Excluded Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Included</th>
<th>Excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>N</td>
<td>1004</td>
<td>649</td>
</tr>
<tr>
<td>Readmitted over</td>
<td>478</td>
<td>321</td>
</tr>
<tr>
<td>180 days (Rate)</td>
<td>(47.6)</td>
<td>(49.5)</td>
</tr>
<tr>
<td>Median age (SD)</td>
<td>62 (15.7)</td>
<td>59 (16.4)</td>
</tr>
<tr>
<td>Females</td>
<td>415</td>
<td>280</td>
</tr>
<tr>
<td></td>
<td>(41.3)</td>
<td>(43.1)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>507</td>
<td>309</td>
</tr>
<tr>
<td></td>
<td>(50.5)</td>
<td>(47.6)</td>
</tr>
<tr>
<td>African American</td>
<td>393</td>
<td>251</td>
</tr>
<tr>
<td></td>
<td>(39.1)</td>
<td>(38.7)</td>
</tr>
<tr>
<td>Other</td>
<td>104</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>(10.4)</td>
<td>(13.7)</td>
</tr>
<tr>
<td>New York Heart Association</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class I</td>
<td>56 (5.6)</td>
<td>48 (7.4)</td>
</tr>
<tr>
<td>Class II</td>
<td>515</td>
<td>309</td>
</tr>
<tr>
<td></td>
<td>(51.3)</td>
<td>(47.6)</td>
</tr>
<tr>
<td>Class III</td>
<td>355</td>
<td>243</td>
</tr>
<tr>
<td></td>
<td>(35.4)</td>
<td>(37.4)</td>
</tr>
<tr>
<td>Class IV</td>
<td>58 (5.8)</td>
<td>41 (6.3)</td>
</tr>
<tr>
<td>Missing</td>
<td>20 (1.9)</td>
<td>8 (1.2)</td>
</tr>
</tbody>
</table>
### Medical History

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVEF(^{†}) % &lt;40</td>
<td>687</td>
<td>448 (68.4) (69.0)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>771</td>
<td>477 (76.8) (73.5)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>450</td>
<td>197 (44.8) (30.3)</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>257</td>
<td>143 (25.6) (22.0)</td>
</tr>
<tr>
<td>Stroke</td>
<td>96 (9.6)</td>
<td>53 (8.2)</td>
</tr>
<tr>
<td>Ischemic Cardiomyopathy</td>
<td>235</td>
<td>141 (23.4) (21.7)</td>
</tr>
</tbody>
</table>

### Clinical Values

*(Mean/SD)*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin</td>
<td>3.32 (0.53)</td>
</tr>
<tr>
<td>Blood Urea</td>
<td>25.2 (17.8)</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>27.1 (18.4)</td>
</tr>
<tr>
<td>Creatinine</td>
<td>1.40 (0.77)</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>12.3 (1.94)</td>
</tr>
<tr>
<td>Glomerular Filtration Rate</td>
<td>58.5 (27.4)</td>
</tr>
<tr>
<td>Potassium</td>
<td>52.8 (31.1)</td>
</tr>
<tr>
<td></td>
<td>4.08 (3.93)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>0.57</td>
<td>0.92</td>
</tr>
</tbody>
</table>

*All values are mean (standard deviation) unless noted.

†LVEF, Left Ventricular Ejection Fraction
Table S6. Percentage of Missing Per Variable from the 236 Variables in the Cohort of 1004 Patients

<table>
<thead>
<tr>
<th>Variable</th>
<th>% Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE_DI_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>AGEOVR90_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>AORTICDISEASE_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>DIABETES_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>ENROLL_LANGUAGE_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>ESSI_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>ETHNICITY_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>FOLSTEIN_SCORE_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>GROUP_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>INHOSPITAL_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>LOW_COG_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>LVEF40_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>MORISKY_CARELESS_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>MORISKY_FORGOT_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>MORISKY_STOP_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>MORISKY_WORSE_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PAYOR_1_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PAYOR_2_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PAYOR_3_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PAYOR_4_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PAYOR_5_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PAYOR_6_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PAYOR_7_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>Variable</td>
<td>Value</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>PAYOR_8_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>PHQ9_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>RACE_1_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>RACE_2_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>RACE_3_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>RACE_4_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>RACE_5_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>RACE_6_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>RACE_7_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>Rx1_0_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>Rx2_0_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>Rx3_0_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>Rx4_0_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>Rx5_0_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>SEX_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>SYMP_BREATH_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>SYMP_CHEST_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>SYMP_FATIGUE_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>SYMP_HEART_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>SYMP_OTHER_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>SYMP_Swell_MISSnIng</td>
<td>0.00</td>
</tr>
<tr>
<td>WARE_1_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>WARE_2_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>WARE_3_MISSING</td>
<td>0.00</td>
</tr>
<tr>
<td>WARE_4_MISSING</td>
<td>0.00</td>
</tr>
</tbody>
</table>
WARE_6_MISSING 0.00
WARE_MISSING 0.00
KCCQ_QUALITY_MISSING 0.10
BMI_MISSING 0.20
ESSI_AFFECTION_MISSING 0.20
KCCQ_EFFICACY_MISSING 0.20
KCCQ_SYMPSCORE_MISSING 0.20
MISSMORISKY_MISSING 0.20
ESSI_CHORES_MISSING 0.30
ESSI_COUNTON_MISSING 0.30
ESSI_ADVICE_MISSING 0.40
KCCQ_HOBBIES_MISSING 0.40
WARE_7_MISSING 0.50
KCCQ_STAIRS_MISSING 0.70
PHQ9_DEPRESSED_MISSING 0.70
WARE_EXPLAIN_MISSING 0.70
MORISKY_MISSING 0.80
PHQ9_CONCENTRATING_MISSING 0.80
KCCQ_WORK_MISSING 0.90
WARE_5_MISSING 0.90
WARE_OFFICE_MISSING 0.90
EDUCATION_MISSING 1.00
ESSI_CLOSE_MISSING 1.00
KCCQ_INTIMATE_MISSING 1.00
KCCQ_SOCIAL_MISSING 1.00
KCCQ_YARDWORK_MISSING  1.00
PHQ9_SLEEPING_MISSING  1.00
AGE_MISSING  1.10
DIFFICULT_MISSING  1.10
INSURANCE_MISSING  1.10
KCCQ_BATHING_MISSING  1.10
KCCQ_BURDEN_MISSING  1.10
KCCQ_SYMPTOMS_MISSING  1.10
PHQ9_SLOW_MISSING  1.10
WARE_CARE_MISSING  1.10
ESSI_EMOTIONAL_MISSING  1.20
KCCQ_DRESSING_MISSING  1.20
KCCQ_VISITING_MISSING  1.20
PHQ9_FAILURE_MISSING  1.20
PHQ9_INTEREST_MISSING  1.20
PHQ9_TIRED_MISSING  1.20
WARE_DOCTOR_MISSING  1.20
WARE_FRIENDLY_MISSING  1.20
MEDCOSTS_MISSING  1.29
WARE_AFFORD_MISSING  1.29
KCCQ_JOGGING_MISSING  1.39
PHQ9_APPETITE_MISSING  1.39
SMOKE_MISSING  1.39
WARE_IGNORE_MISSING  1.39
HEALTH_MISSING  1.49
KCCQ_SHORTNESS_MISSING 1.49
KCCQ_WALKING_MISSING 1.59
WARE_HURRY_MISSING 1.59
KCCQ_DISCOURAGED_MISSING 1.69
KCCQ_FATIGUE_MISSING 1.69
KCCQ_LIMITED_MISSING 1.69
KCCQ_SLEEPING_MISSING 1.69
PHQ9_BOD_MISSING 1.69
WARE_FINANCIAL_MISSING 1.69
WARE_SPECIALIST_MISSING 1.69
WORKPAY_MISSING 1.69
BURDEN_MISSING 1.79
KCCQ_UNDERSTAND_MISSING 1.79
WARE_NEED_MISSING 1.79
HOME_MISSING 1.89
WARE_TIME_MISSING 1.89
AIDS_MISSING 1.99
APNEA_MISSING 1.99
CARDIO_RESYNC_MISSING 1.99
CEREBROVASCULAR_DISEASE_MISSING 1.99
CHRONIC_PULMONARY_MISSING 1.99
CHRONIC_RENAL_FAILURE_MISSING 1.99
CONNECTIVE_TISSUE_MISSING 1.99
CORONARY_ARTERY_MISSING 1.99
DIABETES_ORGAN_MISSING 1.99
DIALYSIS_MISSING 1.99
DIASTOLIC_BP_MISSING 1.99
DIMENTIA_MISSING 1.99
DISEASE_MANAGEMENT_MISSING 1.99
DRUG_USE_MISSING 1.99
HEIGHT_MISSING 1.99
HELP_COMPLETE_MISSING 1.99
HEMIPLEGIA_MISSING 1.99
HIP_MISSING 1.99
HYPERCHOLESTEROLEMIA_MISSING 1.99
HYPERTENSION_MISSING 1.99
ISCHEMIC_CARDIOMYOPATHY_MISSING 1.99
JVD_MISSING 1.99
LEUKEMIA_MISSING 1.99
LIVER_DISEASE_MISSING 1.99
LVEF_METHOD_MISSING 1.99
LYMPHOMA_MISSING 1.99
METASTATIC_TUMOR_MISSING 1.99
NYHA_MISSING 1.99
OXYGEN_MISSING 1.99
PACEMAKER_MISSING 1.99
PATIENT_SCALE_MISSING 1.99
PEPTIC_ULCER_MISSING 1.99
PERIPH_VASCULAR_DISEASE_MISSING 1.99
PITTING_EDEMA_MISSING 1.99
PRIOR_AICD_MISSING  1.99
PRIOR_MI_MISSING  1.99
PRIOR_TIA_MISSING  1.99
PULMONARY_RALES_MISSING  1.99
REALM_ALLERGIC_MISSING  1.99
REALM_ANEMIA_MISSING  1.99
REALM_COLITIS_MISSING  1.99
REALM_CONSTIPATION_MISSING  1.99
REALM_DIRECTED_MISSING  1.99
REALM_FAT_MISSING  1.99
REALM_FATIGUE_MISSING  1.99
REALM_FLU_MISSING  1.99
REALM_JAUNDICE_MISSING  1.99
REALM_NONE_MISSING  1.99
REALM_OSTEOPOROSIS_MISSING  1.99
REALM_PILL_MISSING  1.99
REALM_SCORE_MISSING  1.99
RESP_RATE_MISSING  1.99
RESTING_HR_MISSING  1.99
S3_PRESENCE_MISSING  1.99
SYSTOLIC_BP_MISSING  1.99
TEMP_MISSING  1.99
TUMOR_MISSING  1.99
WAIST_MISSING  1.99
WARE_CAREFUL_MISSING  1.99
<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VISITED_MISSING</td>
<td>4.58</td>
</tr>
<tr>
<td>BLOOD_UREA_NITROGEN_MISSING</td>
<td>5.18</td>
</tr>
<tr>
<td>HEMOCRIT_MISSING</td>
<td>6.27</td>
</tr>
<tr>
<td>PSS_MISSING</td>
<td>7.27</td>
</tr>
<tr>
<td>LVEF_PERCENT_MISSING</td>
<td>7.37</td>
</tr>
<tr>
<td>GFR_MISSING</td>
<td>7.77</td>
</tr>
<tr>
<td>CONTRIBUTE_MISSING</td>
<td>8.57</td>
</tr>
<tr>
<td>HEMOGLOBIN_MISSING</td>
<td>9.46</td>
</tr>
<tr>
<td>KCCQ_FREQUENCY_MISSING</td>
<td>12.85</td>
</tr>
<tr>
<td>DEPENDENT_MISSING</td>
<td>16.63</td>
</tr>
<tr>
<td>KCCQ_SUMMARY_MISSING</td>
<td>17.23</td>
</tr>
<tr>
<td>KNOWLEDGE_MISSING</td>
<td>19.22</td>
</tr>
<tr>
<td>FOLLOWUP_MISSING</td>
<td>19.62</td>
</tr>
<tr>
<td>INCOME_MISSING</td>
<td>19.92</td>
</tr>
<tr>
<td>KCCQ_PHYSICAL_MISSING</td>
<td>21.71</td>
</tr>
<tr>
<td>PATIENT_SCALE_SOURCE_MISSING</td>
<td>24.10</td>
</tr>
<tr>
<td>SCHFI_HEALTH_MISSING</td>
<td>24.50</td>
</tr>
<tr>
<td>SCHFI_BETTER_MISSING</td>
<td>24.80</td>
</tr>
<tr>
<td>SCHFI_SERIOUS_MISSING</td>
<td>24.80</td>
</tr>
<tr>
<td>SCHFI_JUDGE_MISSING</td>
<td>25.30</td>
</tr>
<tr>
<td>BNP_MISSING</td>
<td>26.10</td>
</tr>
<tr>
<td>KCCQ_STABILITY_MISSING</td>
<td>26.99</td>
</tr>
<tr>
<td>DOCYRS_MISSING</td>
<td>35.36</td>
</tr>
<tr>
<td>ALBUMIN_MISSING</td>
<td>42.03</td>
</tr>
<tr>
<td>HOSP_MISSING</td>
<td>48.01</td>
</tr>
<tr>
<td>Variable</td>
<td>Value</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>HF_HOSP_MISSING</td>
<td>53.98</td>
</tr>
<tr>
<td>CHRONIC_RENAL_MISSING</td>
<td>76.69</td>
</tr>
<tr>
<td>VISITEDQT_MISSING</td>
<td>78.88</td>
</tr>
<tr>
<td>JVD_MEASURE_MISSING</td>
<td>87.35</td>
</tr>
<tr>
<td>VISITSPERWEEK_MISSING</td>
<td>87.35</td>
</tr>
<tr>
<td>HF_HOSP_EST_MISSING</td>
<td>87.45</td>
</tr>
<tr>
<td>DOCMOS_MISSING</td>
<td>87.65</td>
</tr>
<tr>
<td>SMOKEQT_MISSING</td>
<td>90.34</td>
</tr>
<tr>
<td>SMOKEAGE_MISSING</td>
<td>90.54</td>
</tr>
<tr>
<td>NT_PRO_BNP_MISSING</td>
<td>91.33</td>
</tr>
<tr>
<td>HEMORRHAGIC_MISSING</td>
<td>93.53</td>
</tr>
<tr>
<td>DOCDAYS_MISSING</td>
<td>94.72</td>
</tr>
<tr>
<td>LIVER_DISEASE_RATING_MISSING</td>
<td>98.01</td>
</tr>
</tbody>
</table>
SUPPLEMENTAL FIGURE

Figure S1

Variable Missing Rates > 2% Missing

<table>
<thead>
<tr>
<th>Variable</th>
<th>% Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIVER_DISEASE_RATING_MISSING</td>
<td>98.01</td>
</tr>
<tr>
<td>DOCDAYS_MISSING</td>
<td>94.72</td>
</tr>
<tr>
<td>HEMORRHAGIC_MISSING</td>
<td>93.53</td>
</tr>
<tr>
<td>NT_PRO_BNP_MISSING</td>
<td>91.33</td>
</tr>
<tr>
<td>SMOKEAGE_MISSING</td>
<td>90.54</td>
</tr>
<tr>
<td>SMOKEQT_MISSING</td>
<td>90.34</td>
</tr>
<tr>
<td>DOCMOS_MISSING</td>
<td>87.65</td>
</tr>
<tr>
<td>HF_HOSP_EST_MISSING</td>
<td>87.45</td>
</tr>
<tr>
<td>VISITSPERWEEK_MISSING</td>
<td>87.35</td>
</tr>
<tr>
<td>JVD_MEASURE_MISSING</td>
<td>87.35</td>
</tr>
<tr>
<td>VISITEDQaising MISSING</td>
<td>78.88</td>
</tr>
<tr>
<td>CHRONIC_RENAL_MISSING</td>
<td>76.69</td>
</tr>
<tr>
<td>HF_HOSP_MISSING</td>
<td>53.98</td>
</tr>
<tr>
<td>HOSP_MISSING</td>
<td>48.01</td>
</tr>
<tr>
<td>ALBUMIN_MISSING</td>
<td>42.03</td>
</tr>
<tr>
<td>DOCYRS_MISSING</td>
<td>35.36</td>
</tr>
<tr>
<td>KCCQ_STABILITY_MISSING</td>
<td>26.99</td>
</tr>
<tr>
<td>BNP_MISSING</td>
<td>26.10</td>
</tr>
<tr>
<td>SCHFI_JUDGE_MISSING</td>
<td>25.30</td>
</tr>
<tr>
<td>SCHFI_SERIOUS_MISSING</td>
<td>24.80</td>
</tr>
<tr>
<td>SCHFI_BETTER_MISSING</td>
<td>24.80</td>
</tr>
<tr>
<td>SCHFI_HEALTH_MISSING</td>
<td>24.50</td>
</tr>
<tr>
<td>PATIENT_SCALE_SOURCE_MISSING</td>
<td>24.10</td>
</tr>
<tr>
<td>KCCQ_PHYSICAL_MISSING</td>
<td>21.71</td>
</tr>
<tr>
<td>INCOME_MISSING</td>
<td>19.92</td>
</tr>
<tr>
<td>FOLLOWUP_MISSING</td>
<td>19.62</td>
</tr>
<tr>
<td>KNOWLEDGE_MISSING</td>
<td>19.22</td>
</tr>
<tr>
<td>KCCQ_SUMMARY_MISSING</td>
<td>17.23</td>
</tr>
<tr>
<td>DEPENDENT_MISSING</td>
<td>16.63</td>
</tr>
<tr>
<td>KCCQ_FREQUENCY_MISSING</td>
<td>12.85</td>
</tr>
<tr>
<td>HEMOGLOBIN_MISSING</td>
<td>9.46</td>
</tr>
<tr>
<td>CONTRIBUTE_MISSING</td>
<td>8.57</td>
</tr>
<tr>
<td>GFR_MISSING</td>
<td>7.77</td>
</tr>
<tr>
<td>LVEF_PERCENT_MISSING</td>
<td>7.37</td>
</tr>
<tr>
<td>PSS_MISSING</td>
<td>7.27</td>
</tr>
<tr>
<td>HEMOCRIT_MISSING</td>
<td>6.27</td>
</tr>
<tr>
<td>BLOOD_UREA_NITROGEN_MISSING</td>
<td>5.18</td>
</tr>
<tr>
<td>VISITED_MISSING</td>
<td>4.58</td>
</tr>
<tr>
<td>POTASSIUM_MISSING</td>
<td>4.48</td>
</tr>
<tr>
<td>FINANCIAL_MISSING</td>
<td>3.98</td>
</tr>
<tr>
<td>ADMITTED_MISSING</td>
<td>3.78</td>
</tr>
<tr>
<td>PSS_DIFFICULTY_MISSING</td>
<td>3.69</td>
</tr>
<tr>
<td>CREATININE_MISSING</td>
<td>3.59</td>
</tr>
</tbody>
</table>