Prediction of Heart Failure Mortality using Machine Learning Algorithms

Rodriguez M, 2021, January.

Abstract

Background

Heart failure continues to present a significant burden on the health system globally. Despite the availability of multiple studies, guidelines, novel medications and technology, the outlook remains poor.

Methodology

A dataset comprising of 13 attributes (including mortality) and 299 instances, collected from admitted heart failure patients in a single institution on 2015 was used. Uni-, bi- and multivariate analyses were performed using sklearn and pandas modules on a Jupyter platform. Linear regression was performed on Ejection Fraction (EF) and age. Modelling was assessed using various attributes on Naive-Bayes, k Nearest Neighbor, Random Forest Classifier and Logistic Regression.

Results

The mean age was 60+/-11 years, and 65% were male. Overall mortality was 32%. Mortality was associated with advanced age, hypertension, decreased EF, anemia, low platelet counts <=100, creatinine >200 mg/dL, low serum sodium <=130 mEq/L, mildly elevated creatinine phosphokinase (CK) (200-500 mcg/L) and markedly elevated CK (>=3000 mcg/L), and short duration of follow-up (< 60 days). Differences in sex, smoking and diabetes status had no effect on mortality. Mortality can be modelled using age, EF and serum creatinine using Naive-Bayes, Random Forest Classifier and Logistic Regression, with an average accuracy score of 75%.

Conclusion

Machine Learning (ML) algorithms can provide acceptable modelling for mortality among heart failure patients. There is a good possibility that utilization of unsupervised ML techniques may provide insights on previously undetected risk factors and relationships.

Keywords: heart failure, heart disease, prediction, mortality, machine learning

Prediction of Heart Failure Mortality using Machine Learning Algorithms

Rodriguez M, 2021, January.

Introduction

Heart Failure (HF) remains a significant contributor for ill health and cause of death globally. It is estimated that 26 million people are affected worldwide.[1] There are 2.6 million Americans affected as of 2016, which was an increase from the previous years, and which is still expected to increase both in number as well as percentage of the population.[2] European trends do not indicate a similar observation of increasing incidence, however, the expectation for HF development is high with a lifetime risk of 28-33% at age 55.[3] It has been observed that less developed countries have even poorer HF prognoses than that of the developed countries in the Americas and Europe.[4]

The causes, manifestations, medical and surgical management of HF are complex. Thus, despite the multiple studies, guidelines and novel therapies implemented, HF continues to pose a challenge to the health care system. This field may exemplify the advantages of Machine Learning (ML) techniques to identify possible correlations that have not been detected by conventional research methods.

The objective of this study is to jumpstart the endeavour towards enabling ML techniques in the predictive analysis for HF. It will start with the basic concepts that deal with categorical and continuous variables, as well as univariate and multivariate analyses. It will then attempt to use the statistical algorithms using ML techniques. This study will provide insights on steps for improvement towards unsupervised ML in helping solve HF dilemmas.

Methodology

The UCI Machine Learning Repository was searched for a reasonably reliable and recent database on heart failure. The Heart failure clinical records Data Set (https://archive.ics.uci.edu/ml/datasets/Heart+failure+clinical+records) was quickly explored using Numbers.

The dataset contained 13 attributes and 299 instances with no missing values. The data was collected from patients admitted in a single institution in 2015. All the patients were aged 40 years and above and were classified to have Left Ventricular Dysfunction (New York Heart Association (NYHA) class III and IV).[5,6]

The attributes collected were as follows:
1. Age (years)
2. Anemia (1= hematocrit <36)
3. high blood pressure (1= present)
4. creatinine phosphokinase (mcg/L)
5. diabetes (1= present)
6. ejection fraction (%)
7. platelets (kiloplatelets/mL)
8. sex (1=male)

9. serum creatinine (mg/dL)
10. serum sodium (Na) (mEq/L)
11. smoking (1= present)
12. follow-up period (days)
13. mortality (1= present)

Several limitations were noted early on, particularly the lack of information regarding the type of HF, the non-definition of smoking and co-morbidities, the lack of hemodynamic measurements, the absence of mention of in-hospital interventions given, and the lack of details regarding hospital discharge. Likewise, the method of data collection and the measures taken to ensure its integrity were not mentioned. Overall, however, the author deemed that the dataset was an acceptable starting point in applying ML analysis to the HF disease spectrum.

To guide the study, the author formulated the following questions:
1. Is higher mortality associated wth advanced age?
2. Is higher mortality affected by the patient's sex?
3. What are the effects of smoking, diabetes and hypertension on mortality?
4. What are the effects of anemia, platelet counts, area, Na and CK on mortality?
5. Is there a higher mortality rate with longer follow-up?

Based on the questions, the following hypotheses were generated:
1. Mortality is associated with advanced age, is higher in males, smokers and patients with DM and hypertension.
2. Anemia, high platelet counts, high area, low Na and high CK increase mortality rates.
3. Mortality rates increase with longer duration of follow-up.

The exploration and analyses were done utilizing the Jupyter Notebook. Modules imported included pandas, numpy, seaboard, matplotlib, statistics and sklearn.

The comma separated values (CSV) form was loaded as a DataFrame, and the dtypes were checked. The absence of null values were confirmed. "High blood pressure" was translated as the presence of "hypertension (HPN)".

Univariate analyses were done using the aggregate and box plot functions. Continuous variables were expressed as frequency, means (or median) +/- standard deviation and range. Binning were done on the continuous variables to be able to extract more meaningful patterns. Categorical variables were expressed as frequencies and percentages.

Bivariate analyses were done primarily using groupby methods, with Mortality as the base attribute (outcome). Percentages were expressed in relation to the group, as well as the whole population (n=299).

Multivariate analyses were analyzed using scatterplot. The choice of the variables were guided by the results from the bivariate analyses.

Forecasting utilized sklearn modules. Linear regression was done to correlate age and EF. Mortality was modelled using Naive-Bayes, k Nearest Neighbors (kNN) (using 5 and 10 neighbors), Random Forest Classifier, and Logistic Regression methods. Initially, all data attributes were utilized as a baseline. The Feature Importances technique was used to identify

the most relevant features for modelling. The metrics module was used to measure the accuracy of the models. Ten accuracy scores were generated and the average was computed for each model. The most important features and most accurate modelling techniques were then chosen to model mortality.

Results

Univariate Analyses

The cohort comprised of 299 patients with a mean age of 60 +/- 11 years (range 40-95). Majority of the cohort was male (65%), and a minority were smokers (32%). Almost half of the patients had diabetes (42%), and 34% had high blood pressure.

The EF of the cohort was low, with a mean of 38 +/- 11% (range 14-80). However, there was a a 10% subset that had an EF of 60% or higher.

Almost half of the cohort was anemic (43%). Majority had normal platelet counts. A small minority had very low counts of <100, or very high counts above 400 (3% and 7%, respectively). Almost 12% of the patients had creatinine levels above 200 mg/dL. Majority had normal serum sodium, however, 7% had levels below 130 mEq/L, and 14% had levels above 140. The serum creatinine phosphokinase levels were unexpectedly relatively low, with a median of 250 mcg/L (range 23- 7861).

The overall duration of follow-up was short, with a mean of 130 +/- 77 days, range (4-285 days). There is a high cohort mortality at 32%.

Bivariate Analyses

There is an increasing trend of mortality rate, starting at 23% for the 40 year olds, up to 83% among the 90 year olds. There is a slightly higher percentage of mortality among hypertensives (37 vs 29%).

The baseline mortality rate is high and continues to worsen as the EF decreases. It ranges from 19% mortality for those with EF of at least 60%, to 80% mortality for EFs less than 20%.

The mortality is slightly higher for anemic patients (36% vs 29%) and patients with low platelet counts <= 100 (62%). Increased creatinine levels are associated with increased mortality. At levels of more than 200 mg/dL, mortality ranges from 43-80%. More than half of the patients (55%) with serum sodium of 130 mEq/L or below resulted in a mortality.

The serum creatinine phosphokinase (CK) showed a biphasic pattern for mortality. When binned for every 100 mcg/L, the mortalities for the 100s', 200's, 300's, 400's and 500's groups (38%, 31%, 41%, 40%, 40%, respectively) were higher than that of levels <100 mcg/L (23%). However, the mortality trends for the levels between 600 and 3000 mcg/L, returns to baseline rates, after which further increase in CK levels are again associated with increased mortality. When binned for every 500, the difference between the 100's and 200's cannot be appreciated, however, the biphasic trend is still apparent (Figure 1).

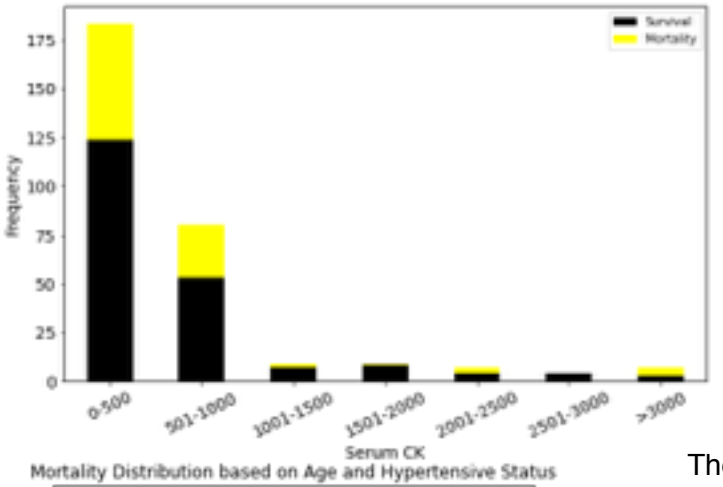


Figure 1. Mortality Distribution of based on creatinine phosphokinase levels

Mortality during the first 60 days of follow-up was high at 83-88%. The mortality rates then stabilize as the follow-up prolongs. This should be evaluated with caution as patients who are counted as mortalities, by default can no longer continue to follow-up.

The outcome did not vary based on sex or status of smoking or diabetes.

<u>Multivariate Analyses</u>

There was a tendency for younger deaths among patients who have HF associated with hypertension and anemia (Figures 2 and 3).

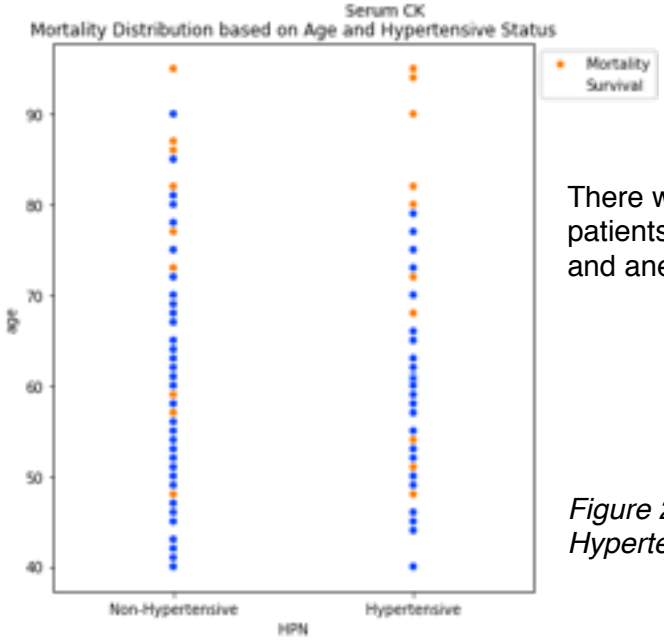


Figure 2. Mortality Distribution based on Age and Hypertension status

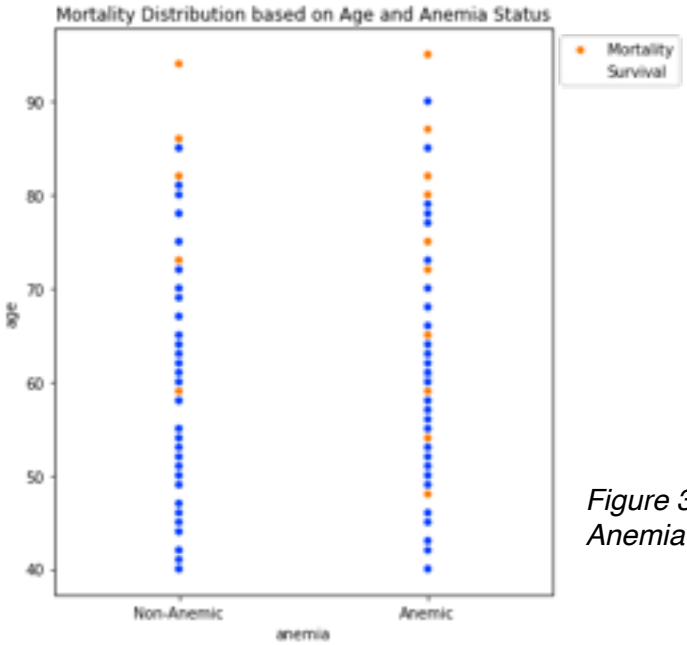
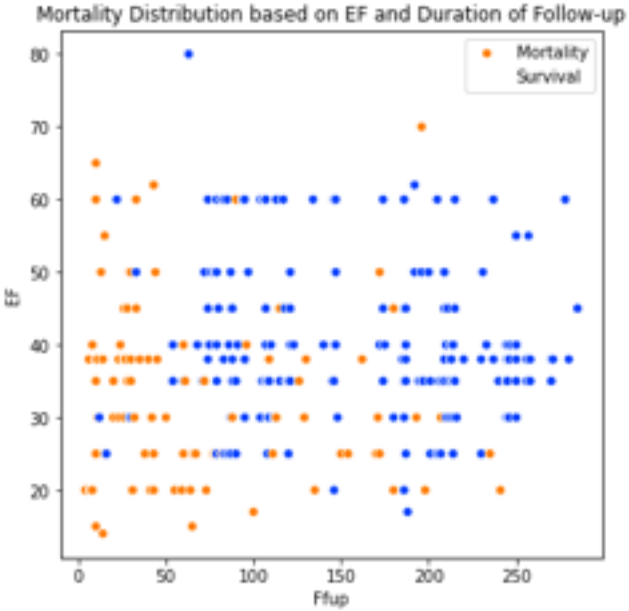


Figure 3. Mortality Distribution based on Age and Anemia status



There was a tendency for mortality on HF patients with low EF and who were non-hypertensive.

Mortalities clustered around the first 60 days of follow-up, especially those with low EF. After the initial 60 days, the survival was mostly sustained for the next 100 days, particularly those with EF of 30 and above (Figure 4).

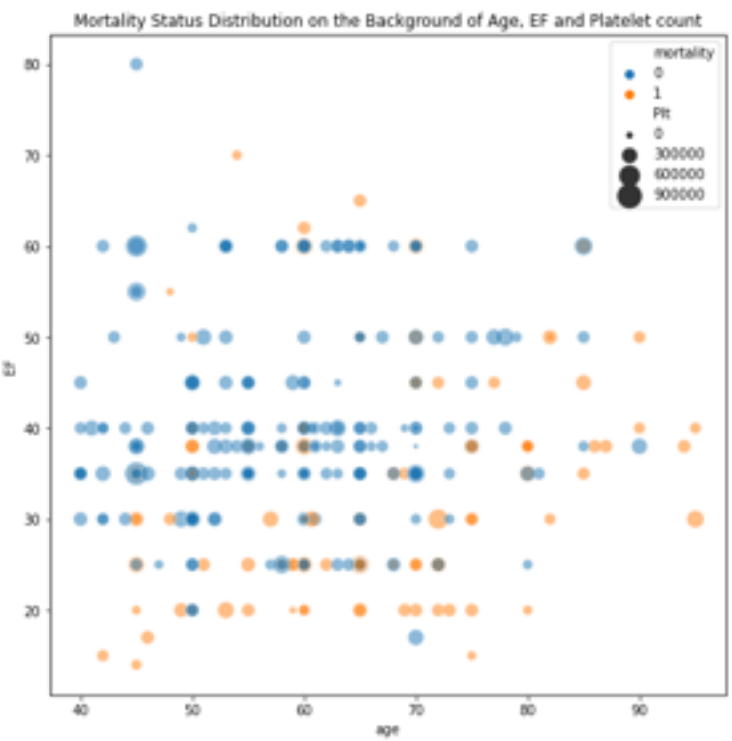
Figure 4. Mortality Distribution based on Ejection Fraction and Duration of Follow-up

Low-EF mortalities were associated with high creatinine levels.

Mortalities associated with very high CK levels were seen as expected. However, the finding of increased mortality in the occurrence of low EF and mildly high CK levels (100-500 mcg/L) lead to a new hypothesis: Are the heart muscles too degenerated to release higher amounts of CK during periods of further stress?

Forecasting

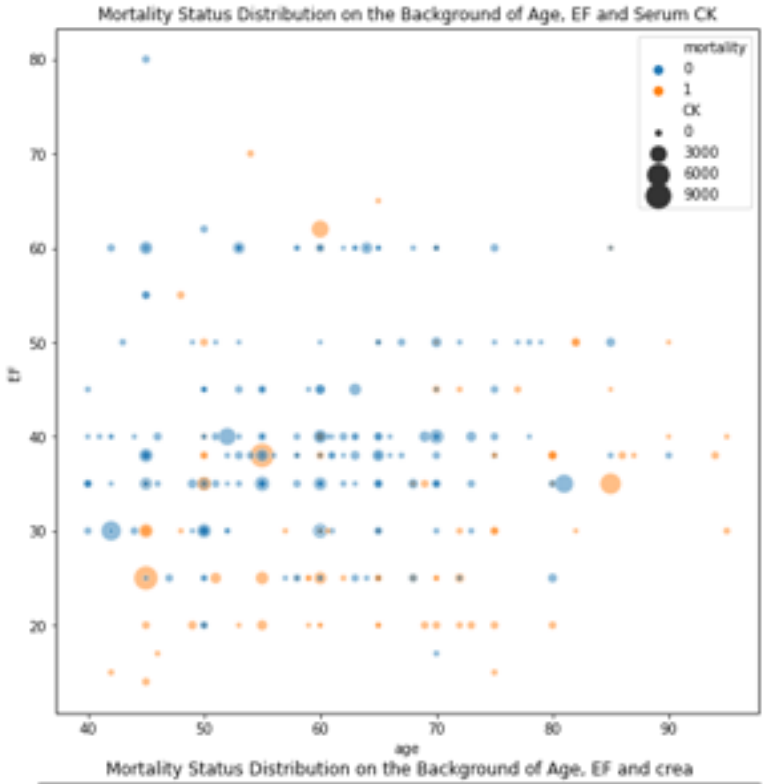
The linear regression model revealed that the deterioration in EF cannot be explained by advancing age.



Using all features for modelling mortality, the following were the average accuracy scores: Naive-Bayes 77%, kNN(5) 62%, kNN(10) 65%, Random Forest Classifier 86%, Logistic Regression 81%.

The feature_importances module showed that the most important features included Ffup, crea, age, EF, CK, Plt and Na, consecutively. Figures 5-7 show the distribution in mortality among the top variables, and help illustrate the modelling.

Figure 5. Mortality Distribution between Age, Ejection Fraction and Platelet count



The top important features were then utilized in re-running the mortality forecasting using the more accurate models. Using the top 6 features (age, EF, Plt, CK, crea and Ffup), for modelling, the average accuracy scores were 75% for Naive-Bayes, 83% for Random Forest Classifier, and 82% for Logistic Regression.

Figure 6. Mortality Distribution between Age, Ejection Fraction and Serum Creatinine Phosphokinase

The Ffup feature was then removed because it was deemed highly susceptible to bias. The average accuracy scores for the 5 retained features were 70-78%.
To see if the model could be simplified without compromising accuracy, only the top three variables were run (age, EF, creatinine). The average accuracy scores were acceptable at 73-76%.

Figure 7. Mortality Distribution between Age, Ejection Fraction and Serum Creatinine

Discussion

Heart Failure is a complex disease and necessitates complex management. Clinicians and researchers have identified multiple variables that contribute to the disease. Multiple studies have led to the creation of risk scores in an attempt to streamline management as well as prognosticate the disease course. The development of the risk scores have utilized different data sources, sample sizes, validation methods, model development and study variables.[7]

Among the identified risk scores with a good C-statistic was the Outcomes of a prospective trial of intravenous milrinone for exacerbation of chronic heart failure (OPTIME HF) score. This study showed that increased age, lower systolic blood pressure, NYHA IV symptoms, elevated blood urea nitrogen (BUN) and decreased sodium during admission are good predictors of mortality within 60 days.[8] Our findings support the OPTIME HF model, with the EF indicating failure symptoms, and creatinine being frequently related to BUN.

The observation that there was a significant proportion of patients with low EF who were not able to release high amounts of CK reflect the theory that failing heart muscles are energy depleted.[9] This suggests that incorporating CK in a scoring system may not be appropriate due to the biphasic pattern.

The 32% mortality rate of this cohort was higher compared to the commonly expected rate of 22% at 1 year.[2] Since this is a cohort from Pakistan, a developing country, the difference in mortality rates could illustrate the importance of socioeconomic variables.

There are available online risk calculators such as the Meta-analysis global group in chronic heart failure (MAGGIC) Risk Calculator and the Seattle Heart Failure Model often require a mix of clinical, laboratory and therapeutic management to compute for expected mortality rates. These help clinicians individualize management for patients and regulate expectancy as well as utilization of resources.

Despite the availability of these risk scoring systems, as well as the advancement in medical science, art and technology, the development and prognosis of heart failure has not improved much through the years.[2,10,11] The risk pictures are changing with an increasing proportion of preserved EF heart failure, as well as a new global outlook where even more variables are seen to contribute significantly to the prognosis, such as economic and social status. This is a good platform to apply techniques on unsupervised machine learning, to uncover unexpected relationships which could be targeted in the hopes of alleviating the burden of disease.

Conclusion

Mortality is associated with advanced age, decreased EF, elevated creatinine levels, decreased platelet counts, decreased sodium levels and anemia. Mortality is apparent only either with mildly high or markedly elevated CK values. Sex, DM, hypertension and smoking status were not associated with mortality. Higher mortality was more evident with shorter follow-up periods. Despite the wide-spread availability of risk scoring for medical management, the challenge of heart failure remains high. The heart failure disease spectrum presents a good field for application of unsupervised machine learning methods to uncover unexpected relationships that could be addressed to decrease the disease burden.

Limitations and Recommendations

The author was not involved in the planning and collection of the dataset. There were limitations in the documentation of the data, as previously mentioned in the methodology.

The model can provide a general overview in predicting mortality for heart failure patients, however, a more robust dataset and analysis will be needed before it can be advocated for clinical use. Important variables specific to HF diagnosis and treatment should be incorporated. The model should be reviewed frequently due to the fast development of medical therapy and technology that revise the clinical trajectories of HF patients.

Significant parallel studies could be done on Quality of Life and Cost-Adjustment parameters.

The HF disease spectrum present a good platform in developing unsupervised ML algorithms.

Disclosure

The author has nothing to disclose.

References

1. Ambrosy A, Fonarow G, Buler J, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. J Am Coll Cardiol. 2014;63:1123-33.

2. Virani S, Alonso A, Benjamin E, et al. Heart Disease and Stroke Statistics - 2020 update: A report from the American Heart Association. Circulation.2020;141:e139-596.

3. Ponikowski P, Voors A, Anker S, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J. 2016;37:2129-200.

4. Dokainish H, Teo K, Zhu J, et al. Global mortality variations in patients with heart failure: Results from the International Congestive Heart Failure (INTER-CHF) Prospective Cohort Study. The Lancet Global Health.2017;5:E665-72.

5. Ahmad T, Munir A, Bhatti S, Aftab M and Raza M. Survival analysis of heart failure patients: A case study. PlosONE. 2017;12:e0181001.

6. Chicco D and Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Medical Informatics and Decision Making. 2020;20:16.

7. Passantino A, Monitillo F, Iacoviello M and Scrutinio D. Predicting mortality in patients with acute heart failure: Role of risk scores. World J Cardiol. 2015;7:902-11.

8. Felker G, Leimberger J, Califf R, et al. Risk stratification after hospitalization for decompensated heart failure. J Card Fail. 2004;10:460-6.

9. Nascimben L, Ingwali J, Paulette P, et al. Creatine kinase system in failing and non failing human moocardium. Circulation. 1996;94:1894-1901.

10. Taylor C, Ordonez-Mena J, Roalfe A, et al. Trends in survival after a diagnosis of heart failure in the United Kingdom 2000-2017: population based cohort study. BMJ.2019;364:l223.

11. Hollenberg S, Stevenson L, Ahmad T, et al. 2019 ACC Expert consensus decision pathway on risk assessment, management, and clinical trajectory of patients hospitalized with heart failure: A report of the American College of Cardiology Solution Set Oversight Committee. J Am Coll Cardiol..2019;74:1966-2011.