



Published in final edited form as:

Heart Fail Clin. 2020 October ; 16(4): 387–407. doi:10.1016/j.hfc.2020.05.002.

Predicting High-Risk Patients and High-Risk Outcomes in Heart Failure

Ramsey M. Wehbe, MD^a, Sadiya S. Khan, MD MSc^{a,b}, Sanjiv J. Shah, MD^a, Faraz S Ahmad, MD MS^{a,b,c}

^aDivision of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

^bDepartment of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

^cCenter for Health Information Partnerships, Institute for Public Health and Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois

Keywords

Heart failure; prognosis; risk factors; risk prediction; risk models; risk scores; artificial intelligence; machine learning; deep learning

“Prediction is very difficult...especially if it’s about the future.”

– Niels Bohr

Introduction

Over 6.5 million Americans are living with heart failure (HF) today and the prevalence continues to increase.¹ Despite substantial advances in medical therapy, many patients progress to end-stage disease with a 5-year absolute mortality rate of approximately 50%.² Moreover, HF is a significant burden on the United States health care system, accounting for ~800,000 hospitalizations in 2016 and projected to cost \$69.7 billion annually by the year 2030.^{1,3}

Identifying high risk HF patients is therefore an important, yet challenging pursuit for clinicians and health care systems alike. Prognosis of individual patients with HF is highly variable in contemporary cohorts, and the risk of serious clinical outcomes such as mortality and hospitalization for HF can differ more than 20-fold.^{4,5} Predicting adverse outcomes in

Corresponding Author: Faraz Ahmad, MD MS. 676 North Saint Clair Street, Suite 600, Chicago, IL 60611. Office: 312-695-4965.

Fax: 312-926-8620. faraz.ahmad@northwestern.edu.

Ramsey M. Wehbe, 676 N St. Clair St, Suite 600, Chicago, IL 60611

Sadiya S. Khan, 676 N St. Clair St, Suite 600, Chicago, IL 60611

Sanjiv J. Shah, 676 N St. Clair St, Suite 600, Chicago, IL 60611

Faraz S. Ahmad, 676 N St. Clair St, Suite 600, Chicago, IL 60611

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

patients with HF could theoretically help direct resources to patients at the highest levels of risk who might benefit the most from earlier and more intensive monitoring and treatment (e.g. targeted medications, cardiac devices, home monitoring systems, and social services), while avoiding unnecessary interventions and costs for patients at low risk.⁶⁻⁹ Ideally, this would translate into improved outcomes and cost-efficiency in providing care for patients with HF. Given the importance of risk prediction for process improvement in HF, we aim to review the available evidence on prognostic variables and the current state of risk prediction for patients with HF. Additionally, we will discuss limitations of traditional risk modeling and provide a glimpse into the future of risk prediction in HF.

Established Risk Factors for Poor Outcomes in Patients with Heart Failure

A number of demographic and clinical variables have been explored as markers of increased risk for adverse outcomes in HF populations. Risk factors can vary substantially based on the outcome of interest or the population under study. Though a discussion of *all* of these risk factors is beyond the scope of this review, some of the prognostic features that have been consistently found to be significant drivers of clinically important outcomes, including mortality, hospitalization, and health-related quality of life (HRQOL) measures are summarized in Table 1. Risk factors are derived from various domains including demographics, clinical characteristics, functional status and HF grade (Table 2), comorbidities, vital signs, labs, imaging, hemodynamics, exercise capacity, medication and device therapy adherence, and social determinants of health.

Multivariable Risk Models For Patients with Heart Failure

Considerable effort has been devoted to developing multivariable risk scores to help summarize and simplify risk assessment so that it can be performed real-time in a clinical setting or embedded into systems of care. To date, hundreds of multivariable risk scores for predicting outcomes in HF populations have been developed. The majority of these risk models have been derived from multivariable statistical modeling such as logistic regression and Cox proportional hazards analysis. Some of these scores are derived directly from parameter estimates of regression models and involve complex calculations, while others have been translated into nomograms with integer scores given to different covariates based on their relative contribution to the overall risk.

As the risk factors that contribute to adverse outcomes can vary substantially based on the HF population (e.g. HFrEF vs. HFpEF) or health care delivery setting (e.g. inpatient vs. outpatient), risk models have been developed for use in specific cohorts including those with chronic ambulatory heart failure, hospitalized heart failure (HHF), and HFpEF specifically. The most popular and well-validated clinical multivariable risk scores for HF are summarized in Table 3, and key features of these risk models are discussed in more detail below. Notably, many of the risk factors discussed previously are shared by several of these multivariable risk models.

The Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement was a consortium document released in 2015 that aimed

to improve the reporting of research surrounding the development and validation of risk scores.^{10,11} The characteristics of model derivation and validation cohorts, outcome of interest, statistical analysis methods, handling of missing data, and measures of performance (e.g. discrimination and calibration) of the model should be clearly defined. Model discrimination, or the ability of a model to distinguish between cases and non-cases, is often measured by the concordance statistic (c-statistic) which is equal to the area under the receiver operator statistic curve (AUC) for a logistic regression model. By convention, an AUC <0.70 indicates inadequate discrimination for clinical use, while an AUC of 0.7–0.8 is considered “acceptable” and an AUC >0.8 is considered “excellent”.¹² Model calibration refers to a model’s ability to accurately predict absolute risk. This is often reported graphically, but there are also several statistical tests of overall calibration which are available.¹³ When available, we will comment on the features discussed above, however, frequently these are not reported for existing heart failure risk models.

Chronic Ambulatory Heart Failure

Numerous risk models have been developed to predict clinical outcomes in cohorts of ambulatory HF patients.^{14–18} The majority of these models have focused on all-cause mortality as the outcome of interest and have examined more intermediate to long term outcomes given short term event rates are typically low in ambulatory populations. Notably, most of the risk models for patients with chronic HF share a core set of common risk factors, including age, sex, NYHA class, BMI, DM, SBP, and indices of renal function.

Perhaps the most popular and thoroughly validated risk model in the chronic ambulatory HF population is the Seattle Heart Failure Model (SHFM).¹⁹ The SHFM was derived from a clinical trial cohort (PRAISE1 trial)²⁰ of 1,125 heart failure patients with a reduced EF (<30%) and NYHA class III or IV symptoms. The components of the risk model were selected via a multivariable stepwise Cox proportional hazards model designed to predict 1, 2, and 3-year mortality. The SHFM was externally validated in 5 separate clinical trial cohorts totaling 9,942 patients from varying populations (HFrEF only, mixed HFrEF/HFpEF, varying functional status and age) with discriminative performance ranging from 0.682 to 0.810 for predicting 1-year mortality and excellent calibration (correlation coefficient = 0.97 between predicted and actual survival for all validation cohorts). The model incorporates 14 continuous variables and 10 categorical variables, including demographics, labs, EKG findings, medications, and device therapy. The authors created an interactive online calculator to facilitate clinical use (<https://depts.washington.edu/shfm>). Since the SHFM was derived from a HF clinical trial cohort almost 30 years ago prior to the widespread use of contemporary GDMT, performance in more modern cohorts, especially at the individual level, has been highly variable and modest at best²¹. Interestingly, the SHFM appears to *overestimate* survival in contemporary HF cohorts, likely explained by the fact that effect sizes for individual medications and ICD/CRT-D use were imputed from results of large randomized clinical trials available at the time.^{18,22} Often, these trials were conducted in a cohort that largely was not on a background of contemporary GDMT, thus likely overestimating individual effect sizes (and consequently overall additive benefit) of these therapies.

More recently, the Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC) developed a heart failure risk score using data on 39,372 patients from 30 cohort studies in an attempt to improve generalizability.²³ The derivation cohorts included patients with HFrEF and HFpEF from both randomized clinical trials and observational studies. Individual components of the model were determined using Poisson regression models with forward stepwise variable selection for the outcome of mortality. Missing data were handled via multiple imputation analysis, a more robust method to handle missing variables than the methods employed in most other risk prediction model derivation strategies. A total of 13 independent predictors were included in the final model, and separate models were derived for patients with HFpEF and HFrEF. Much like the SHFM, the authors provide an online calculator for computing the MAGGIC risk score at the point of care (www.heartfailurerisk.org). Subsequently, the MAGGIC score has been validated in 51,043 patients from a large HF registry cohort (AUC = 0.741 for predicting 3-year mortality)²⁴ and specifically in 407 HFpEF patients from a single institution real-world cohort (AUC = 0.74 for predicting mortality)²⁵. In general, calibration plots in these validation cohorts showed slight underestimation of risk of mortality in high-risk patients and overestimation of risk of mortality in low-risk patients, with better calibration for predicting risk of hospitalization events than the SHFM.

Other notable models that have been developed for predicting outcomes in chronic ambulatory HF include the Candesartan in Heart Failure Assessment of Reduction in Mortality and Morbidity (CHARM)²⁶, Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA)²⁷, and Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico-Heart Failure (GISSI-HF)²⁸ risk models, all of which were derived from clinical trial cohorts. Models derived in clinical trial cohorts are limited by non-representative patient populations due to significant under-recruitment of women and minorities as well as strict inclusion/exclusion criteria and thus have limited generalizability.

Recently, several popular risk scores for chronic ambulatory HF patients were directly compared in a large contemporary European HF registry of over 6,000 patients.¹⁸ The MAGGIC score had the best overall discriminatory power for estimating one-year mortality (AUC 0.743), but all scores were only modest predictors (AUC 0.714–0.743). Moreover, calibration was poor across all risk scores with the MAGGIC, CHARM, and GISSI-HF scores underestimating survival (given they were derived in patients with low background use of contemporary GDMT) and SHFM overestimating survival as explained previously.

Importantly, all of these risk scores were developed prior to the advent of angiotensin-receptor neprilysin inhibitors (ARNIs). A novel risk score was recently published using data from 8,399 patients in the Prospective Comparison of ARNI with ACEI to Determine Impact on Global Mortality and Morbidity in Heart Failure (PARADIGM-HF) trial of sacubitril-valsartan.²⁹ Three separate models were derived for outcomes of cardiovascular death, all-cause mortality, and the composite of cardiovascular death or HF hospitalization (AUC of 0.73, 0.71, and 0.74 at one year, respectively). The models were also validated using data from ~7,000 patients in the Aliskiren Trial to Minimize Outcomes in Patients with Heart Failure (ATMOSPHERE) and ~19,000 patients from the Swedish Heart Failure registry with similar performance. One of the risk factors included in all of the final models

was lack of treatment with ARNI – this risk model may therefore be more appropriate for patients on contemporary GDMT, though given its novelty further validation in prospective cohorts will be necessary to support its clinical use.

Hospitalized Heart Failure

Although there has been progress in improving symptoms and in-hospital mortality for patients hospitalized with HF, events *after* discharge remain unacceptably high with up to 15% of patients dying after discharge and 1 in 4 patients being readmitted within 90 days.^{30,31} Therefore, patients with hospitalized heart failure (HHF) represent a high-risk group for which processes of care designed to improve outcomes with increased resources, intensive monitoring, or targeted interventions may be especially effective. However, even among HHF patients, risk is highly variable. A number of multivariable risk scores have thus been developed specifically to predict subsequent outcomes in patients hospitalized with acute decompensated HF.^{32–37} The majority of these scores focus on short term outcomes of in-hospital mortality or short-term mortality or readmission after discharge. Unlike risk models for chronic ambulatory HF, predictors of outcomes for patients with HHF are more varied, though a few risk factors – namely age, SBP, and renal indices – are shared by many of these models.

In-Hospital Mortality—A few popular clinical risk scores have focused on the outcome of in-hospital mortality, including the Acute Decompensated Heart Failure National Registry (ADHERE) classification and regression tree (CART) model,⁴ Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients With Heart Failure (OPTIMIZE-HF) nomogram,³⁸ and Get With the Guidelines-Heart Failure (GWTG-HF) score.⁵ The ADHERE CART model was derived in a large national registry of ~33,000 patient records with HHF and validated in another ~33,000 patient records using a temporal validation scheme. The resulting ADHERE CART model is a simple tool for risk assessment at the bedside using only 3 predictive variables – blood urea nitrogen (BUN), systolic blood pressure (SBP), and serum creatinine. The discriminatory power of this score was quite modest (AUC 0.67–0.69 in derivation and validation cohorts), however, its allure lies in its simplicity and ease of use at the bedside. The OPTIMIZE-HF nomogram was derived from another national hospital-based registry of almost 50,000 patients in over 250 hospitals. A simplified nomogram for predicting in-hospital mortality was developed using the top 8 predictors from a multivariable logistic regression model and a separate nomogram was later developed for post-discharge outcomes of 60-day mortality and 60-day mortality or HF hospitalization.³⁹ The in-hospital mortality nomogram showed good discrimination with internal validation by bootstrapped resampling (AUC 0.753) as well as in external validation cohorts from the OPTIME-CHF clinical trial (AUC 0.756) and the ADHERE clinical registry (AUC 0.746). Performance was worse, however, for post-discharge nomograms, particularly for the prediction of HF readmission within 60-days. The GWTG-HF Score was derived from a more contemporary registry of approximately 40,000 patients with both HFrEF and HFpEF using a multivariable logistic regression model and internally validated in a subset of the patients from this same registry. Like the ADHERE CART model, age, SBP, and BUN were the strongest contributors to risk of in-hospital mortality, whereas HR, presence of COPD, and serum Na were weaker predictors in the final model. The

model had good discriminatory power in both derivation and validation cohorts (AUC ~ 0.75) regardless of baseline LV systolic function. Calibration plots revealed that the model overestimated risk of mortality in high risk patients.

Post-discharge Outcomes—Though assessing risk of adverse inpatient outcomes may help triage patients who require more immediate attention or intensive monitoring/treatment while hospitalized, post-discharge outcomes remain the greatest source of morbidity in HHF patients. Thus, predicting post-discharge outcomes may be more consequential. Clinical risk prediction models for post-discharge outcomes in patients with HHF have focused on outcomes of all-cause mortality, cardiovascular mortality, all-cause readmission, readmission for HF or cardiovascular causes, or composite endpoints of these outcomes.

The Enhanced Feedback for Effective Cardiac Treatment (EFFECT) and The Evaluation Study of Congestive Heart Failure and Pulmonary Artery Catheterization Effectiveness (ESCAPE) risk scores focused on the outcome of post-discharge mortality. The EFFECT model was derived from a retrospective community-based cohort of 4,031 patients at multiple hospitals in Canada using a multivariable logistic regression model. Predictors of mortality at both 30-days and 1-year were similar, and a composite risk score was able to adequately discriminate very low risk (8% mortality at one year) from high risk (79% mortality at one year) patients with adequate discrimination in the derivation (AUC = 0.77) and internal validation (AUC=0.76) cohorts. The ESCAPE risk score was derived from a North American clinical trial cohort of 433 patients using a Cox proportional hazards model. A simplified risk score for mortality at 6 months was constructed using 8 clinical variables with integer values assigned for each variable's contribution to the overall risk, with BUN and BNP being the most important variables in the model. The model showed good discrimination in derivation (AUC = 0.76) and internal validation with bootstrapping (AUC = 0.78) cohorts, but when externally validated in a clinical trial cohort from the FIRST (Flolan International Randomized Survival Trial) study (without values for BNP or diuretic use available) performed much worse (AUC = 0.65).

In general, models designed to predict non-fatal endpoints such as readmissions perform worse than those models designed to predict mortality. Additionally, predictors of mortality are typically different than those for readmissions. This may reflect the complexity of clinical and non-clinical factors that contribute to risk of hospital readmission and suggests that there may be risk factors that have not yet been identified or that are simply not captured in contemporary data sets. For example, the OPTIME-CHF (Outcomes of a Prospective Trial of Intravenous Milrinone for Exacerbations of Chronic Heart Failure) clinical trial cohort of 949 patients was used to derive two separate models to predict 60-day mortality or the composite of death and rehospitalization at 60 days.⁹ The model for predicting mortality had better discriminatory power (AUC 0.77) than the model for predicting the composite endpoint (AUC 0.69). Similarly, the Placebo-Controlled Randomized Study of the Selective A1 Adenosine Receptor Antagonist Rolofylline for Patients Hospitalized With Acute Decompensated Heart Failure and Volume Overload to Assess Treatment Effect on Congestion and Renal Function (PROTECT) derived risk models for 30-day mortality (AUC 0.79) and 180-day mortality (AUC 0.74) performed better overall than those for predicting 30-day all-cause mortality or hospitalization (AUC 0.66) and 30-day death or

hospitalization for cardiovascular reasons (AUC 0.66).⁴⁰ More recently, the BIOSTAT-CHF project was developed specifically to derive and validate risk prediction models from a large European cohort of 2,516 patients across 69 hospital centers in 11 countries.⁴¹ Though this cohort did not exclusively include patients with HFrEF, the majority of patients were enrolled during an admission for worsening heart failure. Risk models for predicting outcomes of mortality, hospitalization for HF, and the composite outcome of these endpoints were derived from multivariable backward stepwise Cox proportional hazards regression models and validated in an external cohort of 1,738 patients from Scotland. Discriminatory power for the mortality model (AUC 0.73) was remarkably better than for the HF hospitalization model (AUC 0.64) in the external validation cohort and predictors included in respective models were markedly different (Table 3).

Heart Failure with Preserved Ejection Fraction

All of the aforementioned risk models either excluded patients with HFpEF or were derived from mixed populations of HFrEF and HFpEF enrolled in clinical trials, registries, or community cohorts. HFpEF is an increasingly common cause of HF hospitalization in an aging population.⁴² Moreover, the clinical characteristics of patients with HFpEF are distinct – risk models derived from mixed cohorts may fail to capture unique risk factors which contribute to adverse outcomes in this population.^{42–45} Risk scores that have been developed specifically for patients with HFpEF are lacking, though there are a few models that warrant mention.

The I-PRESERVE score was derived from the Irbesartan in Heart Failure with Preserved Ejection Fraction clinical trial cohort of 4,128 patients with *chronic* HFpEF using a forward selection stepwise multivariable Cox proportional hazards model for predicting the outcomes of all-cause mortality, all-cause mortality or HF hospitalization, and HF death or hospitalization.⁴⁶ Interestingly, the variables included in the final derived models are quite similar to those included in the models developed in other chronic HF cohorts discussed previously (Table 3). The model was internally validated with bootstrapped resampling with good discrimination (AUC 0.711–0.765), but was not externally validated and has not been directly compared to other risk scores for chronic HF. In contrast, the Atherosclerosis Risk in Communities (ARIC) study-derived risk score was developed to specifically predict outcomes among patients with HFpEF hospitalized with *acute* decompensated HF.⁴⁷ The authors used a unique approach to build their model using the previously validated EFFECT score (modified to include variables available in the ARIC cohort) as a baseline and evaluating additional candidate variables that further improved prognostication of 28-day or 1-year mortality risk. Interestingly, the final model included several variables that were related to comorbidities implicated in the pathophysiology of the HFpEF syndrome. This model had good discrimination in an internal validation cohort (AUC 0.73 for 28-day and 0.71 for 1-year mortality) with a modest improvement over the EFFECT model alone (AUC 0.70 for 28-day and 0.68 for 1-year mortality). Given the evolving landscape of HF and the epidemiological impact of HFpEF, future research is clearly warranted to further refine risk prediction in this population.

Lack of Evidence for Clinical Utility of Risk Prediction Models

The literature is teeming with HF risk prediction models, however, there is comparatively very little evidence that incorporating risk prediction into clinical practice actually influences management or improves outcomes. Model development for the sake of prediction itself has become a popular “educational hobby”,⁴⁸ but in order for risk models to be used to improve processes of care, they should be *interpretable* and *actionable*. Most models include risk factors that are not modifiable (e.g. age) and it is unclear what, if any, interventions could be implemented to alter prognosis in patients identified as high risk. Additionally, it is unclear what thresholds of risk should be considered “high” or whether there are meaningful thresholds for intervention. Specific interventions to improve adherence to therapies of proven benefit, multidisciplinary disease management programs, and transitional hospital discharge programs have shown promise in improving outcomes (particularly reduction of readmission rates) for patients hospitalized with HF in general,^{49–55} but whether basing these interventions on prognostic information derived from risk models provides any incremental benefit (or reduced costs/resource utilization) is unknown. In those circumstances where prognosis may not be appreciably altered through intervention, there is also limited evidence risk prediction may aid in patient-centered communication regarding risk, preparation for the possibility of death from HF, or referral for advanced HF therapies or interdisciplinary palliative care.⁵⁶

Horne et al. recently published the results of a clinical trial of risk score-guided multidisciplinary team-based care for approximately 6,000 patients hospitalized with HF in 8 hospitals from the Intermountain healthcare system.⁷ Patients who were deemed high risk were identified via an automated risk score calculator embedded in the electronic health record (EHR) and were managed via a distinct care pathway that included personalized changes in their inpatient care, higher-intensity post-discharge follow up, and a more precise discharge plan than lower risk patients. Compared to historical controls, the risk-score guided multidisciplinary care plan resulted in 21% lower 30-day readmission and 52% lower 30-day mortality among high risk patients. Importantly, outcomes among lower risk patients were similar in the intervention and historical control group. Additionally, though inpatient costs were higher in the intervention group, this was balanced by lower post-discharge costs and consequently, there was no difference in overall cost between intervention and control groups. This suggests that the risk-score guided strategy led to improved cost-efficiency in caring for HHF patients.

While these results are promising, this study unfortunately stands alone as the only large clinical trial of its kind. Moreover, reliability of the results may be limited by the use of historical controls rather than randomization, which may have led to confounding (though this was mitigated by the use of a staged crossover design and multivariable analysis that adjusted for potential confounders). Going forward, it is crucial that the value of risk prediction modeling for improving processes of care is rigorously evaluated in randomized clinical trials. The recent high-profile failure of the “hotspotting” program designed reduce spending and improve quality of care among healthcare “superutilizers” underscores the importance of confirming the utility of any presupposed pragmatic intervention in randomized clinical trials before it is adopted into routine practice.⁵⁷ The Risk Evaluation

And Its Impact on Clinical Decision Making and Outcomes in Heart Failure (REVeAL-HF) randomized clinical trial is currently enrolling patients within the Yale New Haven Health System to evaluate the influence of an automatically generated risk prediction score “pop-up” in the EHR on the management and outcomes of HHF patients.⁵⁸ This is hopefully the first of many randomized clinical trials of risk score-based interventions to come.

Barriers to Implementation of Risk Prediction Models in Clinical Practice and Systems of Care

Despite all of the efforts to develop a well-validated risk prediction model for patients with HF, adoption of risk modelling to assess prognosis in clinical settings is quite low. Less than 1% of patients enrolled in a long-term heart failure registry were offered information regarding their prognosis calculated from a risk model by their treating clinician.¹⁸ One major reason for this is likely the lack of evidence for the clinical utility of implementation of HF risk scores, however, there are also several putative barriers to the utilization of heart failure risk predictions scores in practice.

First, the actual performance of these models, especially outside of their derivation cohorts, is disappointing. HF risk models are often derived from single-center populations or clinical trial cohorts without validation in an external cohort. Given the risk factors for adverse outcomes can vary substantially in HF cohorts with different characteristics (e.g. HFrEF vs. HFpEF, clinical trial vs. registry, race/ethnicity, sex, and socioeconomic diversity), the derivation cohort may be very different than the cohort to which the model is applied. Moreover, models derived from historical cohorts may fail to account for the changing landscape of HF and the effect that contemporary therapies have on outcomes - it can be difficult for risk model development to keep pace with therapeutic innovation. Outcomes modeled by different risk scores are often highly variable, including all-cause mortality, CV death, all-cause hospitalizations, HF hospitalizations, and composites of these individual endpoints. Each of these outcomes is associated with a different set of clinical features – indeed, even the 2 major causes of death from HF, sudden cardiac death and progressive pump failure, have distinct risk factors. Additionally, non-fatal events may be more difficult to predict than mortality. We must think carefully about which outcomes are important to predict and prevent – perhaps patient-centered outcomes would be a more appropriate metric. The refractory HF risk score (Table 3) introduced by Allen et al. is unique as it focused on a primary outcome of unfavorable future HRQOL (as assessed by the Kansas City Cardiomyopathy Questionnaire [KCCQ]) or death in the 6 months following admission.⁵⁹ Not only is this a more meaningful metric to patients than readmission rates, it may also identify patients who would benefit from end-of-life discussions or palliative care referral.

Second, HF risk scores rely on multivariable statistical modelling of population data. Selection of predictors is often based on *a priori* knowledge of “traditional” clinical risk factors for the outcome of interest and/or stepwise methods for multivariable selection based on statistical significance, both of which can be problematic and result in significant bias. Predictors included in models are often based on static measurements of clinical variables,

rather than taking advantage of longitudinal dynamic changes in repeated measurements of risk factors which may be informative. Moreover, traditional statistical modeling lacks the ability to model complex, unstructured multi-dimensional data such as raw imaging data or free text from clinical reports. Most available models also fail to indicate how missing data is handled, with even fewer employing robust methods for handling missing data such as multiple imputations. Additionally, popular HF prediction models ignore the presence of competing risks such as from non-cardiac death, heart transplantation, or mechanical circulatory support – an especially salient issue in the advanced heart failure population where competing risks are common.^{60,61} Given statistical models predict outcomes on a population level, it is unsurprising that in practice, these scores perform quite poorly on an individual level. Allen et al. found that popular HF risk models (including SHFM and MAGGIC) performed unacceptably poorly at the individual level, with the majority of patients who died within the next year having greater than a 75% model-predicted probability of survival.²¹

Finally, provider inertia and skepticism present a significant barrier to adoption of risk prediction models in practice, and a majority of providers believe that risk models would not change their practice or add value to their clinical evaluation.^{62,63} Up to three quarters of clinicians surveyed rarely or never use cardiovascular risk prediction scores.⁶² One of the most common reasons provided for not utilizing risk models is the inconvenience and time required to manually calculate risk scores. Though many online calculators exist for popular HF risk scores, these still require manual input of individual risk factors, which can be prohibitively time-consuming especially in the context of increasing pressures on time constraints of the clinical encounter. An effective risk score should therefore be incorporated directly into the EHR with automatic calculation, but there are few current examples of such integration. Providers also express doubts over the utility of risk prediction models, believing these models often oversimplify the risk assessment process and that their own ability to predict risk is superior to that of any model. However, there is evidence that this belief is unfounded. Both providers and patients have proven notoriously bad at predicting prognosis, with providers tending to underestimate survival and patients tending to overestimate survival.^{64–66}

Promise of Artificial Intelligence For Risk Prediction in Heart Failure

Rapid advances in computational power, the digital “big data” revolution, and innovations in mathematical algorithms have led to a recent resurgence of enthusiasm surrounding the utility of artificial intelligence (AI) in a number of industries. The recent popularity of AI has been propelled by the development of novel machine learning (ML) models, which are computer algorithms designed to learn without explicit programming, instead relying on patterns and inference from data (Figure 1). ML has been applied to problems throughout various fields of medicine, particularly radiology, dermatology, and pathology. We are also starting to see the emergence of ML technologies in cardiovascular medicine.

Given the disappointing performance and barriers to use of existing risk scores for HF, there has been substantial interest in novel methods for risk stratification. ML is able to model complex multidimensional interactions within data and ML models are designed

to optimize prediction on an individual level rather than a population level.⁶⁷ Moreover, the *learning* aspect of ML means that these models are modular and adaptable – ML models can be updated to predict outcomes in different populations and more contemporary cohorts as newer data becomes available. Given that ML models are by definition computer programs, they are easily integrated into existing EHR and administrative systems – enabling automated calculation of risk scores towards relieving rather than increasing the burden of resource and time constraints on clinicians and health systems.

Machine Learning for Risk Prediction in Heart failure

Though ML for risk prediction in HF populations is still in its nascence, there are several promising examples of ML-derived risk prediction algorithms that have already been developed (Table 4). Compared to statistically derived risk models, the majority of ML HF risk models incorporate large amounts of features and are derived from very large data sets. Features incorporated into these models include demographic, clinical, physical exam, laboratory, ECG, echocardiographic, medication, procedure, hemodynamic, socioeconomic, quality of life, and administrative claims-based data. Additionally, the majority of these ML-based HF risk models used cross-validation only to assess model performance, with very few using a hold-out validation set and none evaluating model performance in a completely independent test set after training was complete.

Most efforts to develop a ML-derived HF risk model have employed supervised learning algorithms. Several authors have compared HF risk prediction models derived using supervised learning to traditional statistical methodology (namely logistic regression [LR]). Some have found that ML models provided no improvement over LR, while others found modest improvement for ML over LR. Typically, the highest performing ML models are derived from algorithms using ensembles of decision trees, including random forests (RFs) and gradient boosting machines (GBMs). It should be noted, however, that even the best performing ML models showed discriminative power (AUC 0.62 to 0.73) and calibration for prediction of mortality and readmission that is on par with previously reported statistical models in the literature.

Shah et al. utilized a combined approach of unsupervised and supervised learning to identify distinct phenotypes, then predict risk of disease outcomes among a cohort of 397 patients with HFpEF.⁶⁸ Clinical, laboratory, ECG, and echocardiographic data was prospectively collected from 397 patients with HFpEF enrolled in an observational study at a single institution. A total of 46 continuous features (after filtering those that were highly correlated) were utilized in a penalized model-based clustering algorithm. From this analysis 3 clinically distinct phenotypes of HFpEF were identified, each with significantly different risks of mortality or cardiovascular hospitalization. Additionally, using the supervised learning technique of support vector machines (SVMs), a model that included the initial 46 features plus the derived phenogroup feature had fair discriminative ability in predicting the composite outcome of HF hospitalization, cardiovascular hospitalization, or death (AUC 0.67). Ahmad et al. also used a combined approach to develop a risk prediction model among patients in the Swedish Heart Failure Registry by applying supervised learning BEFORE unsupervised learning to model risk.⁶⁹ First, a RFs algorithm was used to identify

predictors of 1-year survival. Discrimination based on the full RFs model (AUC 0.828) was excellent and higher than that historically reported in popular statistical models for predicting 1-year mortality (e.g. SHFM and MAGGIC). Next, unsupervised learning via K-means clustering was applied to the 8 most predictive features identified from the supervised learning algorithm to identify 4 clinically relevant subgroups of HF with distinct risk profiles. Patients in the highest risk cluster had only a 69% one-year survival, compared to 93% one-year survival in the lowest risk group. There are a few other recent examples in the literature of unsupervised learning to identify clusters of patients with distinct risk profiles for clinically meaningful outcomes among patients with chronic heart failure and those hospitalized with acute decompensated heart failure.

Limitations of Machine Learning Risk Models

Thus far, we have yet to see any real adoption of ML risk models for HF in clinical practice or systems of care. ML-derived risk models have generally failed to live up to inflated expectations, showing only modest improvement over traditional statistical models. While some of these shortcomings may simply be due to the fallacy of believing one can predict the future, there are other considerations regarding the potential pitfalls and limitations of ML risk models that should be noted.

Many of the problems that have plagued their statistical predecessors have also troubled ML-derived models. Existing ML models for risk prediction in HF have generally failed to account adequately for time to event outcomes, competing risks, or missing data; though strategies and algorithms to address these problems exist and research in this area is active.^{70,71} Additionally, the majority of ML models utilize discrete values from a single time-point as input features, ignoring the predictive ability of longitudinal data that changes over time. Many ML models have an exceptional ability to analyze time-series data and adjust risk probabilities as new information becomes available, thereby accounting for some of the intrinsically stochastic nature of risk prediction. Further, we should not expect any significant performance boost from simply employing a novel methodology for analyzing data when using the same features/predictors (demographics, vital signs, labs) that were typically used in statistical models⁷². Indeed, one of the strengths of the complexity of ML models, particularly deep learning models, is the ability to handle complex, raw data as inputs as long as the quantity of data is sufficient for the model to learn meaningful representations.⁷³

There are also new considerations presented by ML techniques. Given their complexity, ML models are prone to *overfitting* to training data, which can limit generalizability to external datasets.^{67,73} It is therefore crucial that any ML model is validated multiple times in independent data sets before being adopted for clinical use. Moreover, there are concerns regarding the interpretability of ML models as they can be perceived as being a “black box” due to their focus on the result of the model rather than transparency of the model itself.⁷⁴ Further, ML models often consider thousands of features so it may be difficult to decide which risk factors are most important and consequently, how one could intervene to modify this risk. To address these concerns, new visualization techniques are being developed in order to codify importance of individual features in ML algorithms

(e.g. heatmap visualization of layer class activation in convolutional neural networks) and make these models more actionable.⁷⁵ As with statistically derived risk models, randomized clinical trials assessing the influence of using ML-based HF risk modelling on management and clinically important outcomes are imperative.

Finally, the era of ML brings new cultural and sociological concerns that should be noted.⁷⁴ ML models, particularly supervised learning models, often rely on massive amounts of well-labeled data that requires a significant degree of human input, though there are many strategies to reduce this reliance on human “micro-work” such as semi-supervised learning, data augmentation, transfer learning, fine-tuning, and feature extraction. There is also inherent risk in introducing human implicit bias when developing a ML model, and there are countless examples of unintended and often discriminatory results of ML models.⁷⁶ Lastly, the human-machine interface has become an important consideration and popular culture has promoted a general fear and resistance to the adoption of these technologies that can be perceived to be replacing human input, especially among healthcare providers. However, there are abundant examples in the literature that “augmented intelligence”, that is a human being with the assistance of an artificially intelligent system, almost always outperforms a human or machine alone.⁷²

Summary and Future Directions

In summary, predicting outcomes in patients with HF has proven difficult. While many individual risk factors for adverse future outcomes are known, attempts at developing high-performing risk models in various HF populations using statistically based methods have been generally underwhelming. Moreover, there is currently a lack of evidence to support the notion that implementation of these risk scores into clinical practice or systems of care actually improves outcomes. Newer methods of risk prediction utilizing machine learning technologies offer promise, but thus far have failed to live up to the hype of expectations.

While artificial intelligence may play a role in the future of risk prediction in HF, many key challenges remain. We must think more broadly about the predictors or features included in risk models, including “source” data from unstructured clinical notes text, imaging, pathology slides, hemodynamic and ECG tracings, sensor data (ambulatory monitors, pulmonary artery sensors, and consumer wearables), and other “omics” data (genomics, proteomics, metabolomics, etc.). More robust models can be built using *agnostic* deep learning algorithms, and perhaps the key to predicting the as-of-yet unpredictable is in some of the unrecognized or “hidden” features in this raw data. Additionally, future research is needed to develop better strategies to communicate risk to patients and families in a way that enhances comprehension, identify clinically meaningful thresholds to implement clinical action items (e.g. telemonitoring, pulmonary artery pressure monitoring, multidisciplinary care initiatives), and rigorously evaluate risk-score based strategies in randomized clinical trials. We should strive to constantly assess for bias and ensure that risk models are both equitable and actionable on an individual level – that we are not just predicting the future, but that we are able to intervene and change that future.

Acknowledgments

Disclosure Statement: S.J.S. is supported by grants from the National Institutes of Health (R01 HL140731, R01 HL120728, R01 HL107577, and R01 HL149423); the American Heart Association (#16SFRN28780016, #15CVGSPSD27260148); Actelion, AstraZeneca, Corvia, and Novartis; and has received consulting fees from Actelion, Amgen, AstraZeneca, Bayer, Boehringer-Ingelheim, Cardiora, Eisai, Ionis, Ironwood, Merck, Novartis, Pfizer, Sanofi, and United Therapeutics. F.S.A is supported in part by a grant from the Agency for Healthcare Research and Quality (K12 HS026385) and has received consulting fees from Amgen. The other authors have nothing to disclose.

References

1. Virani SS, Alonso A, Benjamin EJ, et al. Heart Disease and Stroke Statistics—2020 Update A Report From the American Heart Association WRITING GROUP MEMBERS On behalf of the American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. 2020. doi:10.1161/CIR.0000000000000757
2. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: A report of the American college of cardiology foundation/american heart association task force on practice guidelines. *Circulation*. 2013;128(16). doi:10.1161/CIR.0b013e31829e8776
3. Heidenreich PA, Albert NM, Allen LA, et al. Forecasting the impact of heart failure in the united states a policy statement from the american heart association. *Circ Hear Fail*. 2013;6(3):606–619. doi:10.1161/HHF.0b013e318291329a
4. Fonarow GC, Adams KF, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: Classification and regression tree analysis. *J Am Med Assoc*. 2005;293(5):572–580. doi:10.1001/jama.293.5.572
5. Peterson PN, Rumsfeld JS, Liang L, et al. A validated risk score for in-hospital mortality in patients with heart failure from the American heart association get with the guidelines program. *Circ Cardiovasc Qual Outcomes*. 2010;3(1):25–32. doi:10.1161/CIRCOUTCOMES.109.854877 [PubMed: 20123668]
6. O'Connor CM, Hasselblad V, Mehta RH, et al. Triage After Hospitalization With Advanced Heart Failure. The ESCAPE (Evaluation Study of Congestive Heart Failure and Pulmonary Artery Catheterization Effectiveness) Risk Model and Discharge Score. *J Am Coll Cardiol*. 2010;55(9):872–878. doi:10.1016/j.jacc.2009.08.083 [PubMed: 20185037]
7. Horne BD, Roberts CA, Rasmussen KD, et al. Risk score-guided multidisciplinary team-based Care for Heart Failure Inpatients is associated with lower 30-day readmission and lower 30-day mortality. *Am Heart J*. 2020;219:78–88. doi:10.1016/j.ahj.2019.09.004 [PubMed: 31739181]
8. Prasad H, Sra J, Levy WC, Stapleton DD. Influence of predictive modeling in implementing optimal heart failure therapy. *Am J Med Sci*. 2011;341(3):185–190. doi:10.1097/MAJ.0b013e3181ff2393 [PubMed: 21233695]
9. Felker GM, Leimberger JD, Califf RM, et al. Risk stratification after hospitalization for decompensated heart failure. *J Card Fail*. 2004;10(6):460–466. doi:10.1016/j.cardfail.2004.02.011 [PubMed: 15599835]
10. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med*. 2015;162(1):55. doi:10.7326/M14-0697 [PubMed: 25560714]
11. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1–W73. doi:10.7326/M14-0698 [PubMed: 25560730]
12. Lloyd-Jones DM. Cardiovascular risk prediction: Basic concepts, current status, and future directions. *Circulation*. 2010;121(15):1768–1777. doi:10.1161/CIRCULATIONAHA.109.849166 [PubMed: 20404268]
13. Hickey GL, Blackstone EH. External model validation of binary clinical risk prediction models in cardiovascular and thoracic surgery. *J Thorac Cardiovasc Surg*. 2016;152(2):351–355. doi:10.1016/j.jtcvs.2016.04.023 [PubMed: 27215928]

14. Rahimi K, Bennett D, Conrad N, et al. Risk prediction in patients with heart failure: A systematic review and analysis. *JACC Hear Fail.* 2014;2(5):440–446. doi:10.1016/j.jchf.2014.04.008
15. Di Tanna GL, Wirtz H, Burrows KL, Globe G. Evaluating risk prediction models for adults with heart failure: A systematic literature review. *PLoS One.* 2020;15(1):e0224135. doi:10.1371/journal.pone.0224135 [PubMed: 31940350]
16. Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. *JACC Hear Fail.* 2014;2(5):429–436. doi:10.1016/j.jchf.2014.04.006
17. Alba AC, Agoritsas T, Jankowski M, et al. Risk prediction models for mortality in ambulatory patients with heart failure: a systematic review. *Circ Hear Fail.* 2013;6(5):881–889. doi:10.1161/CIRCHEARTFAILURE.112.000043
18. Canepa M, Fonseca C, Chioncel O, et al. Performance of Prognostic Risk Scores in Chronic Heart Failure Patients Enrolled in the European Society of Cardiology Heart Failure Long-Term Registry. *JACC Hear Fail.* 2018;6(6):452–462. doi:10.1016/j.jchf.2018.02.001
19. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: Prediction of survival in heart failure. *Circulation.* 2006;113(11):1424–1433. doi:10.1161/CIRCULATIONAHA.105.584102 [PubMed: 16534009]
20. Packer M, O'Connor CM, Ghali JK, et al. Effect of amlodipine on morbidity and mortality in severe chronic heart failure. *N Engl J Med.* 1996;335(15):1107–1114. doi:10.1056/NEJM199610103351504 [PubMed: 8813041]
21. Allen LA, Matlock DD, Shetterly SM, et al. Use of risk models to predict death in the next year among individual ambulatory patients with heart failure. *JAMA Cardiol.* 2017;2(4):435–441. doi:10.1001/jamacardio.2016.5036 [PubMed: 28002546]
22. Alba AC, Agoritsas T, Jankowski M, et al. Risk prediction models for mortality in ambulatory patients with heart failure: a systematic review. *Circ Hear Fail.* 2013;6(5):881–889. doi:10.1161/CIRCHEARTFAILURE.112.000043
23. Pocock SJ, Ariti CA, McMurray JJV, et al. Predicting survival in heart failure: A risk score based on 39 372 patients from 30 studies. *Eur Heart J.* 2013;34(19):1404–1413. doi:10.1093/eurheartj/ehs337 [PubMed: 23095984]
24. Sartipy U, Dahlström U, Edner M, Lund LH. Predicting survival in heart failure: Validation of the MAGGIC heart failure risk score in 51 043 patients from the Swedish Heart Failure Registry. *Eur J Heart Fail.* 2014;16(2):173–179. doi:10.1111/ejhf.32 [PubMed: 24464911]
25. Rich JD, Burns J, Freed BH, Maurer MS, Burkhoff D, Shah SJ. Meta-analysis global group in chronic (MAGGIC) heart failure risk score: Validation of a simple tool for the prediction of morbidity and mortality in heart failure with preserved ejection fraction. *J Am Heart Assoc.* 2018;7(20):1–12. doi:10.1161/JAHA.118.009594
26. Pocock SJ, Wang D, Pfeffer MA, et al. Predictors of mortality and morbidity in patients with chronic heart failure. *Eur Heart J.* 2006;27(1):65–75. doi:10.1093/eurheartj/ehi555 [PubMed: 16219658]
27. Wedel H, McMurray JJV, Lindberg M, et al. Predictors of fatal and non-fatal outcomes in the Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA): Incremental value of apolipoprotein A-I, high-sensitivity C-reactive peptide and N-terminal pro B-type natriuretic peptide. *Eur J Heart Fail.* 2009;11(3):281–291. doi:10.1093/eurjhf/hfn046 [PubMed: 19168876]
28. Barlera S, Tavazzi L, Franzosi MG, et al. Predictors of mortality in 6975 patients with chronic heart failure in the gruppo italiano per lo studio della streptochinasi nell'infarto miocardico-heart failure trial proposal for a nomogram. *Circ Hear Fail.* 2013;6(1):31–39. doi:10.1161/CIRCHEARTFAILURE.112.967828
29. Simpson J, Jhund PS, Lund LH, et al. Prognostic Models Derived in PARADIGM-HF and Validated in ATMOSPHERE and the Swedish Heart Failure Registry to Predict Mortality and Morbidity in Chronic Heart Failure. *JAMA Cardiol.* 2020;1–10. doi:10.1001/jamacardio.2019.5850
30. Bueno H, Ross JS, Wang Y, et al. Trends in length of stay and short-term outcomes among medicare patients hospitalized for heart failure, 1993–2006. *JAMA - J Am Med Assoc.* 2010;303(21):2141–2147. doi:10.1001/jama.2010.748

31. Gheorghide M, Shah AN, Vaduganathan M, et al. Recognizing Hospitalized Heart Failure as an Entity and Developing New Therapies to Improve Outcomes. Academics', Clinicians', Industry's, Regulators', and Payers' Perspectives. *Heart Fail Clin.* 2013;9(3):285–290. doi:10.1016/j.hfc.2013.05.002 [PubMed: 23809415]
32. Ross JS, Mulvey GK, Stauffer B, et al. Statistical models and patient predictors of readmission for heart failure: A systematic review. *Arch Intern Med.* 2008;168(13):1371–1386. doi:10.1001/archinte.168.13.1371 [PubMed: 18625917]
33. Fonarow GC. Clinical Risk Prediction Tools in Patients Hospitalized with Heart Failure. *Rev Cardiovasc Med.* 2012;13(1):14–23. doi:10.3909/ricm0595
34. Passantino A. Predicting mortality in patients with acute heart failure: Role of risk scores. *World J Cardiol.* 2015;7(12):902. doi:10.4330/wjc.v7.i12.902 [PubMed: 26730296]
35. Saito M, Negishi K, Marwick TH. Meta-Analysis of Risks for Short-Term Readmission in Patients with Heart Failure. *Am J Cardiol.* 2016;117(4):626–632. doi:10.1016/j.amjcard.2015.11.048 [PubMed: 26772444]
36. Lagu T, Pekow PS, Shieh MS, et al. Validation and comparison of seven mortality prediction models for hospitalized patients with acute decompensated heart failure. *Circ Hear Fail.* 2016;9(8):1–10. doi:10.1161/CIRCHEARTFAILURE.115.002912
37. Mahajan SM, Heidenreich P, Abbott B, Newton A, Ward D. Predictive models for identifying risk of readmission after index hospitalization for heart failure: A systematic review. *Eur J Cardiovasc Nurs.* 2018. doi:10.1177/1474515118799059
38. Abraham WT, Fonarow GC, Albert NM, et al. Predictors of In-Hospital Mortality in Patients Hospitalized for Heart Failure. Insights From the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients With Heart Failure (OPTIMIZE-HF). *J Am Coll Cardiol.* 2008;52(5):347–356. doi:10.1016/j.jacc.2008.04.028 [PubMed: 18652942]
39. O'Connor CM, Abraham WT, Albert NM, et al. Predictors of mortality after discharge in patients hospitalized with heart failure: An analysis from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). *Am Heart J.* 2008;156(4):662–673. doi:10.1016/j.ahj.2008.04.030 [PubMed: 18926148]
40. Cleland JG, Chiswell K, Teerlink JR, et al. Predictors of postdischarge outcomes from information acquired shortly after admission for acute heart failure: A report from the placebo-controlled randomized study of the selective α_1 adenosine receptor antagonist rolofylline for patients hospitalized w. *Circ Hear Fail.* 2014;7(1):76–87. doi:10.1161/CIRCHEARTFAILURE.113.000284
41. Voors AA, Ouwerkerk W, Zannad F, et al. Development and validation of multivariable models to predict mortality and hospitalization in patients with heart failure. *Eur J Heart Fail.* 2017;19(5):627–634. doi:10.1002/ejhf.785 [PubMed: 28247565]
42. Steinberg BA, Zhao X, Heidenreich PA, et al. Trends in patients hospitalized with heart failure and preserved left ventricular ejection fraction: Prevalence, therapies, and outcomes. *Circulation.* 2012;126(1):65–75. doi:10.1161/CIRCULATIONAHA.111.080770 [PubMed: 22615345]
43. Vaduganathan M, Fonarow GC. Epidemiology of Hospitalized Heart Failure. Differences and Similarities Between Patients with Reduced versus Preserved Ejection Fraction. *Heart Fail Clin.* 2013;9(3):271–276. doi:10.1016/j.hfc.2013.04.001 [PubMed: 23809413]
44. Doughty RN, Cubbon R, Ezekowitz J, et al. The survival of patients with heart failure with preserved or reduced left ventricular ejection fraction: An individual patient data meta-analysis: Meta-analysis Global Group in Chronic Heart Failure (MAGGIC). *Eur Heart J.* 2013;33(14):1750–1757. doi:10.1093/eurheartj/ehr254
45. Kapoor JR, Kapoor R, Ju C, et al. Precipitating Clinical Factors, Heart Failure Characterization, and Outcomes in Patients Hospitalized With Heart Failure With Reduced, Borderline, and Preserved Ejection Fraction. *JACC Hear Fail.* 2016;4(6):464–472. doi:10.1016/j.jchf.2016.02.017
46. Komajda M, Carson PE, Hetzel S, et al. Factors associated with outcome in heart failure with preserved ejection fraction. *Circ Hear Fail.* 2011;4(1):27–35. doi:10.1161/CIRCHEARTFAILURE.109.932996
47. Thorvaldsen T, Claggett BL, Shah A, et al. Predicting Risk in Patients Hospitalized for Acute Decompensated Heart Failure and Preserved Ejection Fraction: The Atherosclerosis Risk in

- Communities Study Heart Failure Community Surveillance. *Circ Hear Fail*. 2017;10(12):1–10. doi:10.1161/CIRCHEARTFAILURE.117.003992
48. Stevenson LW, Davis RB. Model building as an educational hobby. *Circ Hear Fail*. 2016;9(8). doi:10.1161/CIRCHEARTFAILURE.116.003457
 49. Ahmad FS, Metlay JP, Barg FK, Henderson RR, Werner RM. Identifying Hospital Organizational Strategies to Reduce Readmissions. *Am J Med Qual*. 2013;28(4):278–285. doi:10.1177/1062860612464999 [PubMed: 23118202]
 50. McAlister FA, Stewart S, Ferrua S, McMurray JJV. Multidisciplinary strategies for the management of heart failure patients at high risk for admission: A systematic review of randomized trials. *J Am Coll Cardiol*. 2004;44(4):810–819. doi:10.1016/j.jacc.2004.05.055 [PubMed: 15312864]
 51. Holland R, Battersby J, Harvey I, Lenaghan E, Smith J, Hay L. Systematic review of multidisciplinary interventions in heart failure. *Heart*. 2005;91(7):899–906. doi:10.1136/hrt.2004.048389 [PubMed: 15958358]
 52. Murray MD, Young J, Hoke S, et al. Pharmacist intervention to improve medication adherence in heart failure: A randomized trial. *Ann Intern Med*. 2007;146(10):714–725. doi:10.7326/0003-4819-146-10-200705150-00005 [PubMed: 17502632]
 53. Granger BB, Ekman I, Hernandez AF, et al. Results of the chronic heart failure intervention to improve medication adherence study: A randomized intervention in high-risk patients. *Am Heart J*. 2015;169(4):539–548. doi:10.1016/j.ahj.2015.01.006 [PubMed: 25819861]
 54. McAlister FA, Lawson FME, Teo KK, Armstrong PW. A systematic review of randomized trials of disease management programs in heart failure. *Am J Med*. 2001;110(5):378–384. doi:10.1016/S0002-9343(00)00743-9 [PubMed: 11286953]
 55. Peiris D, Usherwood T, Panaretto K, et al. Effect of a computer-guided, quality improvement program for cardiovascular disease risk management in primary health care: The treatment of cardiovascular risk using electronic decision support cluster-randomized trial. *Circ Cardiovasc Qual Outcomes*. 2015;8(1):87–95. doi:10.1161/CIRCOUTCOMES.114.001235 [PubMed: 25587090]
 56. Rogers JG, Patel CB, Mentz RJ, et al. Palliative Care in Heart Failure: The PAL-HF Randomized, Controlled Clinical Trial. *J Am Coll Cardiol*. 2017. doi:10.1016/j.jacc.2017.05.030
 57. Finkelstein A, Zhou A, Taubman S, Doyle J. Health care hotspotting - A randomized, controlled trial. *N Engl J Med*. 2020;382(2):152–162. doi:10.1056/NEJMsa1906848 [PubMed: 31914242]
 58. Tariq Ahmad MM, Nihar R Desai MM, Francis P Wilson MM. Risk Evaluation And Its Impact on ClinicAl Decision Making and Outcomes in Heart Failure (REVeAL-HF). [ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT03845660](https://clinicaltrials.gov/ct2/show/NCT03845660). Accessed February 5, 2020.
 59. Allen LA, Gheorghiade M, Reid KJ, et al. Identifying patients hospitalized with heart failure at risk for unfavorable future quality of life. *Circ Cardiovasc Qual Outcomes*. 2011;4(4):389–398. doi:10.1161/CIRCOUTCOMES.110.958009 [PubMed: 21693723]
 60. Fang J, Abdel-Qadir H, Austin PC, et al. Importance of Considering Competing Risks in Time-to-Event Analyses. *Circ Cardiovasc Qual Outcomes*. 2018;11(7). doi:10.1161/circoutcomes.118.004580
 61. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation*. 2016;133(6):601–609. doi:10.1161/CIRCULATIONAHA.115.017719 [PubMed: 26858290]
 62. Eichler K, Zoller M, Tschudi P, Steurer J. Barriers to apply cardiovascular prediction rules in primary care: A postal survey. *BMC Fam Pract*. 2007;8. doi:10.1186/1471-2296-8-1 [PubMed: 17346340]
 63. Sposito AC, Ramires JAF, Wouter Jukema J, et al. Physicians' attitudes and adherence to use of risk scores for primary prevention of cardiovascular disease: Cross-sectional survey in three world regions. *Curr Med Res Opin*. 2009;25(5):1171–1178. doi:10.1185/03007990902846423 [PubMed: 19323611]
 64. Yamokoski LM, Hasselblad V, Moser DK, et al. Prediction of Rehospitalization and Death in Severe Heart Failure by Physicians and Nurses of the ESCAPE Trial. *J Card Fail*. 2007;13(1):8–13. doi:10.1016/j.cardfail.2006.10.002 [PubMed: 17338997]

65. Allen LA, Yager JE, Funk MJ, et al. Discordance between patient-predicted and model-predicted life expectancy among ambulatory patients with heart failure. *JAMA - J Am Med Assoc.* 2008;299(21):2533–2542. doi:10.1001/jama.299.21.2533
66. Buchan TA, Ross HJ, McDonald M, et al. Physician judgment versus model predicted prognosis in patients with heart failure. *Can J Cardiol.* 2019;36(1):84–91. doi:10.1016/j.cjca.2019.07.623 [PubMed: 31735429]
67. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med.* 2016;375(13):1216–1219. doi:10.1056/NEJMp1606181 [PubMed: 27682033]
68. Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation.* 2015;131(3):269–279. doi:10.1161/CIRCULATIONAHA.114.010637 [PubMed: 25398313]
69. Ahmad T, Pencina MJ, Schulte PJ, et al. Clinical implications of chronic heart failure phenotypes defined by cluster analysis. *J Am Coll Cardiol.* 2014;64(17):1765–1774. doi:10.1016/j.jacc.2014.07.979 [PubMed: 25443696]
70. Krittanawong C, Johnson KW, Rosenson RS, et al. Deep learning for cardiovascular medicine: A practical primer. *Eur Heart J.* 2019;40(25):2058–2069C. doi:10.1093/eurheartj/ehz056 [PubMed: 30815669]
71. Wang P, Li Y, Reddy CK. Machine learning for survival analysis: A survey. *ACM Comput Surv.* 2019;51(6). doi:10.1145/3214306
72. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. *N Engl J Med.* 2017;376(26):2507–2509. doi:10.1056/NEJMp1702071 [PubMed: 28657867]
73. Deo RC. Machine learning in medicine. *Circulation.* 2015;132(20):1920–1930. doi:10.1161/CIRCULATIONAHA.115.001593 [PubMed: 26572668]
74. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician humanism and artificial intelligence. *JAMA - J Am Med Assoc.* 2018;319(1):19–20. doi:10.1001/jama.2017.19198
75. Samek W, Binder A, Montavon G, Lapuschkin S, Müller KR. Evaluating the visualization of what a deep neural network has learned. *IEEE Trans Neural Networks Learn Syst.* 2017;28(11):2660–2673. doi:10.1109/TNNLS.2016.2599820
76. Hajian S, Bonchi F, Castillo C. Algorithmic bias: From discrimination discovery to fairness-aware data mining. *Proc ACM SIGKDD Int Conf Knowl Discov Data Min.* 2016;13–17-Aug:2125–2126. doi:10.1145/2939672.2945386
77. Ho KKL, Anderson KM, Kannel WB, Grossman W, Levy D. Survival after the onset of congestive heart failure in Framingham Heart Study subjects. *Circulation.* 1993;88(1):107–115. doi:10.1161/01.CIR.88.1.107 [PubMed: 8319323]
78. Wong CM, Hawkins NM, Jhund PS, et al. Clinical characteristics and outcomes of young and very young adults with heart failure: The CHARM programme (candesartan in heart failure assessment of reduction in mortality and morbidity). *J Am Coll Cardiol.* 2013;62(20):1845–1854. doi:10.1016/j.jacc.2013.05.072 [PubMed: 23850914]
79. Romiti GF, Recchia F, Zito A, Visioli G, Basili S, Raparelli V. Sex and Gender-Related Issues in Heart Failure. *Heart Fail Clin.* 2020;16(1):121–130. doi:10.1016/j.hfc.2019.08.005 [PubMed: 31735310]
80. Glynn P, Lloyd-Jones DM, Feinstein MJ, Carnethon M, Khan SS. Disparities in Cardiovascular Mortality Related to Heart Failure in the United States. *J Am Coll Cardiol.* 2019;73(18):2354–2355. doi:10.1016/j.jacc.2019.02.042 [PubMed: 31072580]
81. Akintoye E, Briasoulis A, Egbe A, et al. National trends in admission and in-hospital mortality of patients with heart failure in the United States (2001–2014). *J Am Heart Assoc.* 2017;6(12). doi:10.1161/JAHA.117.006955
82. Massimo P, Egidio T, Giovanni C, Rachel S, Maurizio S, Luigi T. Loading Manipulations Improve the Prognostic Value of Doppler Evaluation of Mitral Flow in Patients With Chronic Heart Failure. *Circulation.* 1997;95(5):1222–1230. doi:10.1161/01.CIR.95.5.1222 [PubMed: 9054853]

83. Felker GM, Thompson RE, Hare JM, et al. Underlying causes and long-term survival in patients with initially unexplained cardiomyopathy. *N Engl J Med.* 2000;342(15):1077–1084. doi:10.1056/NEJM200004133421502 [PubMed: 10760308]
84. Taylor MRG, Fain PR, Sinagra G, et al. Natural history of dilated cardiomyopathy due to lamin A/C gene mutations. *J Am Coll Cardiol.* 2003;41(5):771–780. doi:10.1016/S0735-1097(02)02954-6 [PubMed: 12628721]
85. The Criteria Committee of the New York Heart Association. Classification of Functional Capacity and Objective Assessment. In: *Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels.* 9th ed. Little, Brown & Co; 1994:253–256.
86. Goldman L, Hashimoto B, Cook EF, Loscalzo A. Comparative reproducibility and validity of systems for assessing cardiovascular functional class: Advantages of a new specific activity scale. *Circulation.* 1981;64(6):1227–1234. doi:10.1161/01.CIR.64.6.1227 [PubMed: 7296795]
87. Taichman DB, McGoon MD, Harhay MO, et al. Wide variation in clinicians' assessment of New York Heart Association/World Health Organization functional class in patients with pulmonary arterial hypertension. *Mayo Clin Proc.* 2009;84(7):586–592. doi:10.4065/84.7.586 [PubMed: 19567712]
88. Hunt SA, Abraham WT, Chin MH, et al. 2009 Focused Update Incorporated Into the ACC/AHA 2005 Guidelines for the Diagnosis and Management of Heart Failure in Adults. A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines Developed. *J Am Coll Cardiol.* 2009;53(15). doi:10.1016/j.jacc.2008.11.013
89. Stevenson LW, Paganì FD, Young JB, et al. INTERMACS Profiles of Advanced Heart Failure: The Current Picture. *J Hear Lung Transplant.* 2009;28(6):535–541. doi:10.1016/j.healun.2009.02.015
90. Valentova M, Anker SD, von Haehling S. Cardiac Cachexia Revisited: The Role of Wasting in Heart Failure. *Heart Fail Clin.* 2020;16(1):61–69. doi:10.1016/j.hfc.2019.08.006 [PubMed: 31735316]
91. Shindler DM, Kostis JB, Yusuf S, et al. Diabetes mellitus, a predictor of morbidity and mortality in the Studies of Left Ventricular Dysfunction (SOLVD) trials and registry. *Am J Cardiol.* 1996;77(11):1017–1020. doi:10.1016/S0002-9149(97)89163-1 [PubMed: 8644628]
92. Schmidt M, Ulrichsen SP, Pedersen L, Bøtker HE, Sørensen HT. Thirty-year trends in heart failure hospitalization and mortality rates and the prognostic impact of co-morbidity: A Danish nationwide cohort study. *Eur J Heart Fail.* 2016;18(5):490–499. doi:10.1002/ejhf.486 [PubMed: 26868921]
93. Lawson CA, Mamas MA, Jones PW, et al. Association of Medication Intensity and Stages of Airflow Limitation With the Risk of Hospitalization or Death in Patients With Heart Failure and Chronic Obstructive Pulmonary Disease. *JAMA Netw open.* 2018;1(8):e185489. doi:10.1001/jamanetworkopen.2018.5489 [PubMed: 30646293]
94. Xanthopoulos A, Starling RC, Kitai T, Triposkiadis F. Heart Failure and Liver Disease: Cardiohepatic Interactions. *JACC Hear Fail.* 2019;7(2):87–97. doi:10.1016/j.jchf.2018.10.007
95. Costanzo MR. The Cardiorenal Syndrome in Heart Failure. *Heart Fail Clin.* 2020;16(1):81–97. doi:10.1016/j.hfc.2019.08.010 [PubMed: 31735318]
96. Lam CSP, Teng THK. Minding the gap in heart failure. Understanding the pulse pressure in reduced versus preserved ejection fraction. *JACC Hear Fail.* 2016;4(1):50–54. doi:10.1016/j.jchf.2015.10.005
97. Chin MH, Goldman L. Correlates of early hospital readmission or death in patients with congestive heart failure. *Am J Cardiol.* 1997;79(12):1640–1644. doi:10.1016/S0002-9149(97)00214-2 [PubMed: 9202355]
98. Chahal H, Bluemke DA, Wu CO, et al. Heart failure risk prediction in the Multi-Ethnic Study of Atherosclerosis. *Heart.* 2015;101(1):58–64. doi:10.1136/heartjnl-2014-305697 [PubMed: 25381326]
99. Lee TT, Chen J, Cohen DJ, Tsao L. The association between blood pressure and mortality in patients with heart failure. *Am Heart J.* 2006;151(1):76–83. doi:10.1016/j.ahj.2005.03.009 [PubMed: 16368295]
100. Fox K, Ford I, Steg PG, Tendera M, Robertson M, Ferrari R. Heart rate as a prognostic risk factor in patients with coronary artery disease and left-ventricular systolic dysfunction (BEAUTIFUL):

- a subgroup analysis of a randomised controlled trial. *Lancet*. 2008;372(9641):817–821. doi:10.1016/S0140-6736(08)61171-X [PubMed: 18757091]
101. Obeidat M, Burgess M, Lip GYH. Atrial Fibrillation in Heart Failure: Focus on Antithrombotic Management. *Heart Fail Clin*. 2020;16(1):107–120. doi:10.1016/j.hfc.2019.08.007 [PubMed: 31735308]
 102. Linssen GCM, Rienstra M, Jaarsma T, et al. Clinical and prognostic effects of atrial fibrillation in heart failure patients with reduced and preserved left ventricular ejection fraction. *Eur J Heart Fail*. 2011;13(10):1111–1120. doi:10.1093/eurjhf/hfr066 [PubMed: 21642293]
 103. Santhanakrishnan R, Wang N, Larson MG, et al. Atrial fibrillation begets heart failure and vice versa: Temporal associations and differences in preserved versus reduced ejection fraction. *Circulation*. 2016;133(5):484–492. doi:10.1161/ARCTICULATIONAHA.U5.018614 [PubMed: 26746177]
 104. Wang NC, Maggioni AP, Konstam MA, et al. Clinical implications of QRS duration in patients hospitalized with worsening heart failure and reduced left ventricular ejection fraction. *JAMA - J Am Med Assoc*. 2008;299(22):2656–2666. doi:10.1001/jama.299.22.2656
 105. Ellis ER, Josephson ME. Ventricular Arrhythmias and Heart Failure. In: Eisen H, ed. *Heart Failure: A Comprehensive Guide to Pathophysiology and Clinical Care*. London: Springer London; 2017:339–369. doi:10.1007/978-1-4471-4219-5_15
 106. Lee WH, Packer M. Prognostic importance of serum sodium concentration and its modification by converting-enzyme inhibition in patients with severe chronic heart failure. *Circulation*. 1986;73(2):257–267. doi:10.1161/01.CIR.73.2.257 [PubMed: 3002660]
 107. Fonarow GC, Peacock WF, Phillips CO, Givertz MM, Lopatin M. Admission B-Type Natriuretic Peptide Levels and In-Hospital Mortality in Acute Decompensated Heart Failure. *J Am Coll Cardiol*. 2007;49(19):1943–1950. doi:10.1016/j.jacc.2007.02.037 [PubMed: 17498579]
 108. Kociol RD, Horton JR, Fonarow GC, et al. Admission, discharge, or change in B-type natriuretic peptide and long-term outcomes: Data from Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF) linked to Medicare claims. *Circ Heart Fail*. 2011;4(5):628–636. doi:10.1161/CIRCHEARTFAILURE.111.962290
 109. Savarese G, Musella F, D'Amore C, et al. Changes of natriuretic peptides predict hospital admissions in patients with chronic heart failure. A meta-analysis. *JACC Hear Fail*. 2014;2(2):148–158. doi:10.1016/j.jchf.2013.11.007
 110. Khanam SS, Son JW, Lee JW, et al. Prognostic value of short-term follow-up BNP in hospitalized patients with heart failure. *BMC Cardiovasc Disord*. 2017;17(1):1–10. doi:10.1186/s12872-017-0632-0 [PubMed: 28052754]
 111. Richter B, Koller L, Hohensinner PJ, et al. A multi-biomarker risk score improves prediction of long-term mortality in patients with advanced heart failure. *Int J Cardiol*. 2013;168(2):1251–1257. doi:10.1016/j.ijcard.2012.11.052 [PubMed: 23218577]
 112. Carluccio E, Dini FL, Biagioli P, et al. The “Echo Heart Failure Score”: An echocardiographic risk prediction score of mortality in systolic heart failure. *Eur J Heart Fail*. 2013;15(8):868–876. doi:10.1093/eurjhf/hft038 [PubMed: 23512095]
 113. Hamada-Harimura Y, Seo Y, Ishizu T, et al. Incremental Prognostic Value of Right Ventricular Strain in Patients With Acute Decompensated Heart Failure. *Circ Cardiovasc Imaging*. 2018;11(10):e007249. doi:10.1161/ARCIIMAGING.U7.007249 [PubMed: 30354477]
 114. Park JJ, Park JB, Park JH, Cho GY. Global Longitudinal Strain to Predict Mortality in Patients With Acute Heart Failure. *J Am Coll Cardiol*. 2018;71(18):1947–1957. doi:10.1016/j.jacc.2018.02.064 [PubMed: 29724346]
 115. Marwick TH, Shah SJ, Thomas JD. Myocardial Strain in the Assessment of Patients with Heart Failure: A Review. *JAMA Cardiol*. 2019;4(3):287–294. doi:10.1001/jamacardio.2019.0052 [PubMed: 30810702]
 116. Kuruvilla S, Adenaw N, Katwal AB, Lipinski MJ, Kramer CM, Salerno M. Late gadolinium enhancement on cardiac magnetic resonance predicts adverse cardiovascular outcomes in nonischemic cardiomyopathy: A systematic review and meta-analysis. *Circ Cardiovasc Imaging*. 2014;7(2):250–257. doi:10.1161/CIRCIMAGING.113.001144 [PubMed: 24363358]

117. Halliday BP, Baksi AJ, Gulati A, et al. Outcome in Dilated Cardiomyopathy Related to the Extent, Location, and Pattern of Late Gadolinium Enhancement. *JACC Cardiovasc Imaging*. 2019;12(8P2):1645–1655. doi:10.1016/j.jcmg.2018.07.015 [PubMed: 30219397]
118. Craig M, Pereira NL. Right heart catheterization and risk stratification in advanced heart failure. *Curr Heart Fail Rep*. 2006;3(3):143–152. doi:10.1007/s11897-006-0014-x [PubMed: 16914107]
119. Cooper LB, Mentz RJ, Stevens SR, et al. Hemodynamic Predictors of Heart Failure Morbidity and Mortality: Fluid or Flow? *J Card Fail*. 2016;22(3):182–189. doi:10.1016/j.cardfail.2015.11.012 [PubMed: 26703245]
120. Fonarow Gregg. Treatment Targets in Heart Failure continued Hemodynamics in Decompensated Heart Failure. *Rev Cardiovasc Med*. 2001;2(suppl 2):s7–s12. http://medreviews.com/sites/default/files/2016-11/RICM_2suppl2_s7.pdf. [PubMed: 12439356]
121. Nohria A, Tsang SW, Fang JC, et al. Clinical assessment identifies hemodynamic profiles that predict outcomes in patients admitted with heart failure. *J Am Coll Cardiol*. 2003;41(10):1797–1804. doi:10.1016/S0735-1097(03)00309-7 [PubMed: 12767667]
122. Ghio S, Gavazzi A, Campana C, et al. Independent and additive prognostic value of right ventricular systolic function and pulmonary artery pressure in patients with chronic heart failure. *J Am Coll Cardiol*. 2001;37(1):183–188. doi:10.1016/S0735-1097(00)01102-5 [PubMed: 11153735]
123. Patel CB, Devore AD, Felker GM, et al. Characteristics and outcomes of patients with heart failure and discordant findings by right-sided heart catheterization and cardiopulmonary exercise testing. *Am J Cardiol*. 2014;114(7):1059–1064. doi:10.1016/j.amjcard.2014.07.022 [PubMed: 25212547]
124. Borlaug BA, Nishimura RA, Sorajja P, Lam CSP, Redfield MM. Exercise hemodynamics enhance diagnosis of early heart failure with preserved ejection fraction. *Circ Hear Fail*. 2010;3(5):588–595. doi:10.1161/CIRCHEARTFAILURE.109.930701
125. Methvin A, Georgiopolou VV., Kalogeropoulos AP, et al. Usefulness of Cardiac Index and Peak Exercise Oxygen Consumption for Determining Priority for Cardiac Transplantation. *Am J Cardiol*. 2010;105(9):1353–1355. doi:10.1016/j.amjcard.2009.12.053 [PubMed: 20403492]
126. Myers J, Wong M, Adhikarla C, et al. Cardiopulmonary and noninvasive hemodynamic responses to exercise predict outcomes in heart failure. *J Card Fail*. 2013;19(2):101–107. doi:10.1016/j.cardfail.2012.11.010 [PubMed: 23384635]
127. Keteyian SJ, Patel M, Kraus WE, et al. Variables measured during cardiopulmonary exercise testing as predictors of mortality in chronic systolic heart failure. *J Am Coll Cardiol*. 2016;67(7):780–789. doi:10.1016/j.jacc.2015.11.050 [PubMed: 26892413]
128. Setoguchi S, Stevenson LW, Schneeweiss S. Repeated hospitalizations predict mortality in the community population with heart failure. *Am Heart J*. 2007;154(2):260–266. doi:10.1016/j.ahj.2007.01.041 [PubMed: 17643574]
129. Krumholz HM, Chen YT, Wang Y, Vaccarino V, Radford MJ, Horwitz RI. Predictors of readmission among elderly survivors of admission with heart failure. *Am Heart J*. 2000;139(1 D):72–77. doi:10.1016/S0002-8703(00)90311-9 [PubMed: 10618565]
130. Arundel C, Lam PH, Khosla R, et al. Association of 30-Day All-Cause Readmission with Long-Term Outcomes in Hospitalized Older Medicare Beneficiaries with Heart Failure. *Am J Med*. 2016;129(11):1178–1184. doi:10.1016/j.amjmed.2016.06.018 [PubMed: 27401949]
131. Greene SJ, Butler J, Albert NM, et al. Medical Therapy for Heart Failure With Reduced Ejection Fraction: The CHAMP-HF Registry. *J Am Coll Cardiol*. 2018;72(4):351–366. doi:10.1016/j.jacc.2018.04.070 [PubMed: 30025570]
132. Greene SJ, Fonarow GC, DeVore AD, et al. Titration of Medical Therapy for Heart Failure With Reduced Ejection Fraction. *J Am Coll Cardiol*. 2019;73(19):2365–2383. doi:10.1016/j.jacc.2019.02.015 [PubMed: 30844480]
133. Roth GA, Poole JE, Zaha R, Zhou W, Skinner J, Morden NE. Use of guideline-directed medications for heart failure before cardioverter-defibrillator implantation. *J Am Coll Cardiol*. 2016;67(9):1062–1069. doi:10.1016/j.jacc.2015.12.046 [PubMed: 26940927]
134. Joynt Maddox KE, Reidhead M, Hu J, et al. Adjusting for social risk factors impacts performance and penalties in the hospital readmissions reduction program. *Health Serv Res*. 2019;54(2):327–336. doi:10.1111/1475-6773.13133 [PubMed: 30848491]

135. Meddings J, Reichert H, Smith SN, et al. The Impact of Disability and Social Determinants of Health on Condition-Specific Readmissions beyond Medicare Risk Adjustments: A Cohort Study. *J Gen Intern Med.* 2017;32(1):71–80. doi:10.1007/s11606-016-3869-x [PubMed: 27848189]
136. Foraker RE, Rose KM, Suchindran CM, Chang PP, McNeill AM, Rosamond WD. Socioeconomic status, medicaid coverage, clinical comorbidity, and rehospitalization or death after an incident heart failure hospitalization atherosclerosis risk in communities cohort (1987 to 2004). *Circ Hear Fail.* 2011;4(3):308–316. doi:10.1161/CIRCHEARTFAILURE.110.959031
137. Bikkeli B, Wayda B, Bao H, et al. Place of residence and outcomes of patients with heart failure: Analysis from the telemonitoring to improve heart failure outcomes trial. *Circ Cardiovasc Qual Outcomes.* 2014;7(5):749–756. doi:10.1161/CIRCOUTCOMES.113.000911 [PubMed: 25074375]
138. Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting Mortality among Patients Hospitalized for Heart Failure: Derivation and Validation of a Clinical Model. *J Am Med Assoc.* 2003;290(19):2581–2587. doi:10.1001/jama.290.19.2581
139. Loghmanpour NA, Kormos RL, Kanwar MK, Teuteberg JJ, Murali S, Antaki JF. A Bayesian Model to Predict Right Ventricular Failure Following Left Ventricular Assist Device Therapy. *JACC Hear Fail.* 2016;4(9):711–721. doi:10.1016/j.jchf.2016.04.004
140. Mortazavi BJ, Downing NS, Buchholz EM, et al. Analysis of Machine Learning Techniques for Heart Failure Readmissions. *Circ Cardiovasc Qual Outcomes.* 2016;9(6):629–640. doi:10.1161/CIRCOUTCOMES.116.003039 [PubMed: 28263938]
141. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: Comparison of machine learning and other statistical approaches. *JAMA Cardiol.* 2017;2(2):204–209. doi:10.1001/jamacardio.2016.3956 [PubMed: 27784047]
142. Ahmad T, Lund LH, Rao P, et al. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J Am Heart Assoc.* 2018;7(8). doi:10.1161/JAHA.117.008081
143. Golas SB, Shibahara T, Agboola S, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: A retrospective analysis of electronic medical records data. *BMC Med Inform Decis Mak.* 2018;18(1). doi:10.1186/s12911-018-0620-z
144. Adler ED, Voors AA, Klein L, et al. Improving risk prediction in heart failure using machine learning. *Eur J Heart Fail.* 2019;1–9. doi:10.1002/ejhf.1628 [PubMed: 30644642]
145. Angraal S, Mortazavi BJ, Gupta A, et al. Machine Learning Prediction of Mortality and Hospitalization in Heart Failure with Preserved Ejection Fraction. *JACC Hear Fail.* 2019;8(1). doi:10.1016/j.jchf.2019.06.013
146. Desai RJ, Wang SV, Vaduganathan M, Evers T, Schneeweiss S. Comparison of Machine Learning Methods With Traditional Models for Use of Administrative Claims With Electronic Medical Records to Predict Heart Failure Outcomes. *JAMA Netw Open.* 2020;3(1):1–15. doi:10.1001/jamanetworkopen.2019.18962

Synopsis:

Identifying patients with heart failure at high risk for poor outcomes is important for patient care, resource allocation, and process improvement. Although numerous risk models exist to predict mortality, hospitalization, and patient-reported health status, they are infrequently used for several reasons, including modest performance, lack of evidence to support routine clinical use, and barriers to implementation. Artificial intelligence has the potential to enhance the performance of risk prediction models, but has its own limitations and remains unproven.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Key Points:

- Identifying patients with heart failure who are at high risk for poor outcomes is important for patient care, resource utilization, and process improvement
- Several risk models exist to identify high risk patients with heart failure, based on traditional statistical risk modeling methodology
- Risk models for heart failure are infrequently utilized in practice due to their modest performance outside validation cohorts, lack of evidence to support that risk modeling in heart failure improves outcomes, and barriers to implementation
- Machine learning may offer an alternative solution for a precision medicine approach to personalized risk prediction in heart failure patients, but has its own limitations
- The future of risk prediction in HF will likely involve more sophisticated risk modeling, incorporating diverse data sources not typically included in existing risk scores

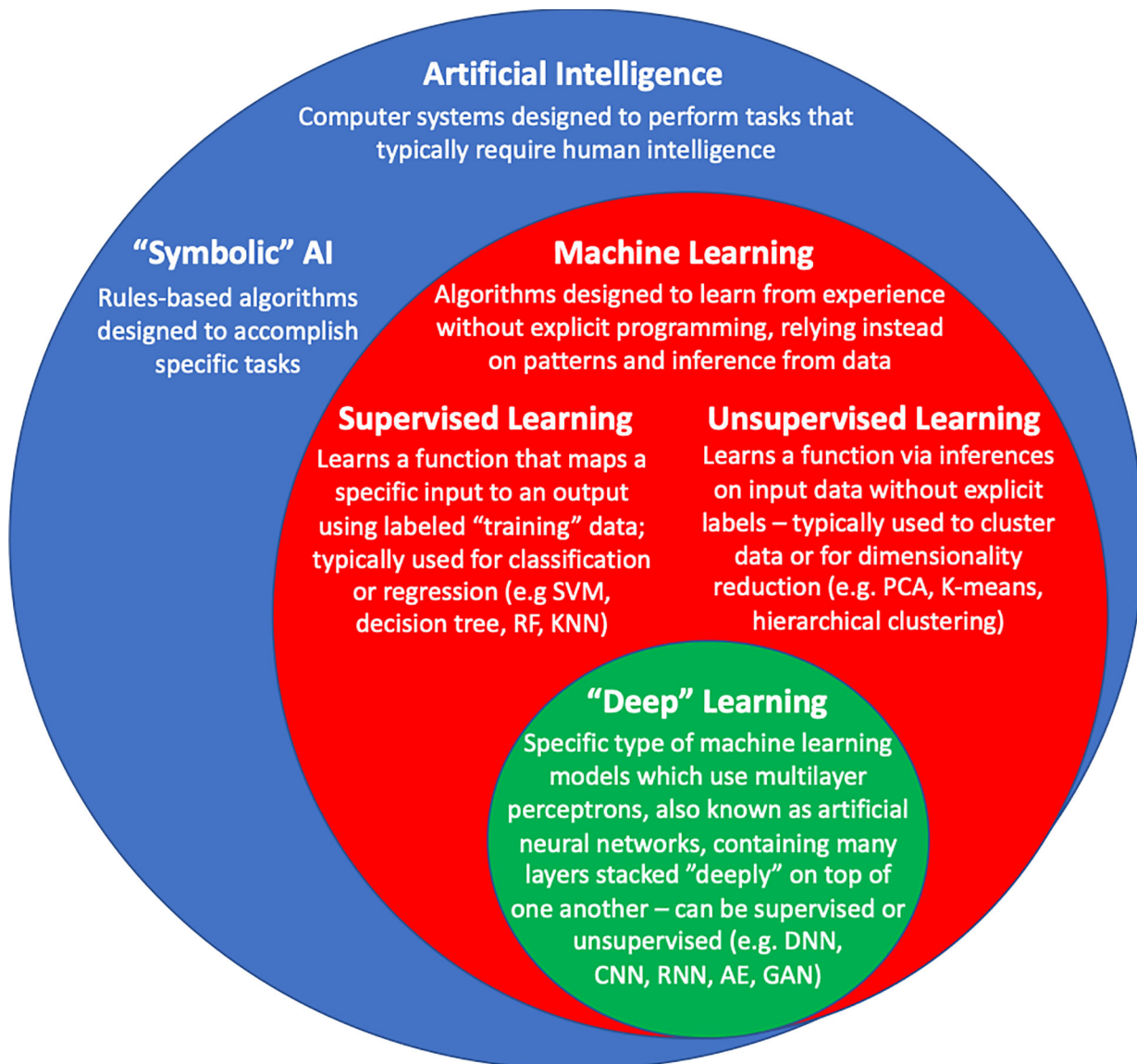


Figure 1. Artificial Intelligence and Machine Learning. Schema of the relationship between artificial intelligence, machine learning, and “deep” learning. SVM = Support vector machines, RF = random forests, KNN = k-nearest neighbor, PCA = principal components analysis, DNN = deep neural network, CNN = convolutional neural network, RNN = recurrent neural network, GAN = generative adversarial network, VAE = variational autoencoder

Table 1.

Established Risk Factors for Poor Prognosis in Patients with Heart Failure

Risk Factors
Demographics ^{14,15,77-81} – Age, sex, race/ethnicity
Systolic and diastolic dysfunction ^{42-46,82}
Heart failure etiology ^{83,84} (e.g. ischemic vs.non-ischemic)
Functional class, ⁸⁵⁻⁸⁷ heart failure stage, ⁸⁸ INTERMACS profile ⁸⁹
Body Mass Index ⁹⁰ – Cardiac cachexia and the “obesity paradox”
Comorbidities ^{2,26,91-95} – Diabetes, renal dysfunction, liver disease, COPD, anemia, depression
Vital signs ⁹⁶⁻¹⁰⁰ – Hypotension, HR, HR variability
Electrophysiologic risk factors ¹⁰¹⁻¹⁰⁵ – Arrhythmias (AF, VT/VF), QRS prolongation
Laboratory values ¹⁰⁶⁻¹¹¹ – Sodium, biomarkers (natriuretic peptides, troponin, novel biomarkers - ST-2, galectin-3, GDF-15)
Imaging ¹¹²⁻¹¹⁷ – Chamber enlargement, right ventricular dysfunction, strain imaging, delayed gadolinium enhancement
Hemodynamics ¹¹⁸⁻¹²⁴ – Elevated filling pressures, pulmonary hypertension, cardiac index, exercise hemodynamics
Exercise capacity ¹²⁵⁻¹²⁷ – Peak VO ₂ , 6MWT
Prior hospitalization for HF ¹²⁸⁻¹³⁰
Underutilization of guideline directed medical therapies ¹³¹⁻¹³³
Social determinants of health ¹³⁴⁻¹³⁷

Table 2.

ACCF/AHA HF Stage, NYHA Functional Class, and INTERMACS Profiles^{85,88,89}

ACCF/AHA Stages	NYHA Class	INTERMACS Profile
A – At risk for HF, but no structural heart disease		
B – Structural heart disease, but no symptoms or limitations	I – Asymptomatic, no physical limitations	
C – Structural heart disease with current or prior heart failure symptoms	II – Symptomatic with moderate exertion, slight limitation in activity	
	III – Symptomatic with minimal exertion, marked limitation in activity	7- Advanced NYHA III 6- Exertion limited
D – End-stage disease, refractory heart failure requiring specialized interventions	IV – Symptomatic at rest, inability to exert oneself	5- Exertion intolerant
		4- Resting symptoms
		3- Inotrope dependent, but stable
		2- Progressive decline
		1- Critical cardiogenic shock

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Popular Clinical Multivariable Risk Prediction Models for Patients with Heart Failure

Risk Model	Multivariate Predictors	Outcomes
Chronic HF		
<i>Seattle Heart Failure Model</i> ¹⁹ (2006)	Age, Sex, LVEF, NYHA Class, Weight, SBP, Labs (Hgb, Lymphs, UA, Chol, Na), Wide QRS/LBBB, Meds (including GDMT, allopurinol, statins, diuretics), Devices (ICD/CRT/PPM)	1, 2, and 5-year mortality
<i>CHARM Risk Score</i> ²⁶ (2006)	Age, Sex, LVEF, NYHA Class, BMI, DM, AF, MR, MI, Prior HF hosp, HF duration, Current smoker, DBP, HR, Dependent edema, Pulmonary crackles, Dyspnea, Pulmonary edema, Cardiomegaly, ARB	All-cause mortality; CV death or HF hosp
<i>CORONA Risk Score</i> ²⁷ (2009)	Age, Sex, LVEF, NYHA Class, BMI, DM, MI, CABG, AF, Claudication, HR, NT-proBNP, Cr, ApoA-I	CV death, non-fatal MI, or stroke
<i>GISSI-HF Model</i> ²⁸ (2013)	Age, Sex, LVEF, NYHA Class, BMI, DM, COPD, AS, SBP, Labs (eGFR, Hgb, UA, NT-proBNP, hs-CTnT)	All-cause mortality
<i>MAGGIC Risk Score</i> ²³ (2013)	Age, Sex, LVEF, NYHA class, BMI, DM, COPD, Recent diagnosis, Smoking status, SBP, Cr, Meds (B-blockers, ACEi/ARB)	1 and 3-year all-cause mortality
<i>PARADIGM Risk Score</i> ²⁹ (2020)	Age, Sex, Race, LVEF, NYHA Class, BMI, DM, Region, HF duration, HF hosp, MI, Valve dz, BBB, PCI, PAD, SBP, Bili, AST, UA, Alb, K, Cl, Hgb, LDL, Trig, BUN, WBCs, NT-proBNP, B-blocker, ARNI	CV Death or HF Hosp; CV death; All-cause death
Hospitalized HF		
<i>EFFECT Model</i> ¹³⁸ (2003)	Age, COPD, Cirrhosis, Cancer, Dementia, Cerebrovascular dz, SBP, RR, Na, BUN, Hgb	30-day and 1-year mortality
<i>OPTIME-CHF Model</i> ⁹ (2004)	Age, NYHA Class, SBP, BUN, Na Prior HF hosp, SBP, BUN, h/o PCI	60-day mortality 60-day readmission
<i>ADHERE CART Model</i> ⁴ (2005)	SBP, BUN, Creatinine	In-hospital mortality
<i>OPTIMIZE-HF Nomogram</i> ^{38,39} (2008)	Age, LVEF, HF primary diagnosis, HR, SBP, Na, Cr Age, Weight, SBP, Na, Cr, Liver dz, Depression, Reactive airway dz	In-hospital mortality 60-day mortality or readmission
<i>GWTG-HF Score</i> ⁵ (2010)	Age, Race, COPD, SBP, HR, Na, BUN	In-hospital mortality
<i>ESCAPE Risk Score</i> ⁶ (2010)	Age, CPR/mechanical ventilation, Na, BNP, BUN, Diuretic, B-Blocker, 6MWT	6-month mortality
<i>Refractory HF Risk Score</i> ⁵⁹ (2011)	Age, DM, Stroke, Arrhythmia, Na, BNP, BUN, Betablocker, KCCQ Score	Unfavorable QOL or death at 6 months
<i>PROTECT HF Post-discharge Model</i> ⁴⁰ (2014)	Age, SBP, Na, Cr, BUN, Albumin, Prior HF Hosp, Peripheral edema	30/180-day mortality, All-cause Hosp, and/or CV Hosp
<i>BIOSTAT-CHF Model</i> ⁴¹ (2017)	Age, BUN, NT-proBNP, Hgb, B-blocker Age, prior HF hosp, edema, SBP, eGFR	All-cause mortality HF hosp
HFpEF		
<i>I-PRESERVE Score</i> ⁴⁶ (2011)	Age, LVEF, Ischemic etiology, Prior HF hosp, QOL, DM, MI, COPD/asthma, HR, eGFR, Neutrophils, NT-proBNP	All-cause mortality or CV hosp, all-cause mortality, HF hosp or mortality
<i>ARIC Score</i> ⁴⁷ (2017)	Modified EFFECT score + Race, BMI, AF, HR, Hypoxia, Natriuretic peptides	28-day and 1-year mortality

Table 4.

Machine Learning Models For Predicting Outcomes in Patients with HF

Study	Population	Target (Outcomes)	Features (Predictors)	Models	Performance	Validation/Test sets
Shah et al. (2015) ⁶⁸	Recently hospitalized outpatients with HFpEF – N = 420	CV Hosp, HF Hosp, and Death	46 features - clinical variables, physical characteristics, labs, ECG, and echo parameters	Phenotyping via model-based clustering + Risk prediction via SVM	Combined model AUC 0.67 for primary outcome	Hold-out validation -107 patients with HFpEF
Loghmanpour et al. (2016) ¹³⁹	Pre-LVAD patients – N = 10,909	Right ventricular failure after LVAD	33 features - pre-op demographics, clinical features, labs, echo, hemodynamics, others	Tree-augmented naïve Bayesian network	AUC 0.903 for acute RVF (vs. Drakos 0.547 and RVFRS 0.498)	10-fold cross-validation
Mortazavi et al. (2016) ¹⁴⁰	Recently discharged HF patients – N = 1,655	30 and 180-day all-cause readmission and readmission for HF	236 features – demographics, labs, physical exam, quality of life, SES	RF, GBM, RF/SVM, RF/LR vs. LR	All-cause hosp - RF (best) (AUC 0.628 vs. 0.533 LR) HF hosp – GBM (best) (AUC 0.678 vs. 0.543 LR)	50% random hold-out validation
Frizzel et al. (2017) ¹⁴¹	Inpatients with HF from GW/ TG-HF registry – N = 56,477	30-day all cause readmission	86 features - demographics, SES, medical history, characterization of HF; meds, vital signs, weights, labs, treatment, and discharge interventions	Tree-augmented naïve Bayesian network, LR with LASSO, RF, and GBM	AUC 0.607–0.618 vs. 0.624 for LR	30% random hold-out validation
Ahmad et al. (2018) ¹⁴²	Swedish HF registry–N = 44,886	1-year survival	86 features - demographics, clinical, laboratory, and imaging data	RF, K-means clustering (using top 8 features from RF)	RF AUC 0.83, 8 cluster variables AUC 0.78	10-fold cross validation
Golas et al. (2018) ¹⁴³	Patients with a history of HF admission – N = 11,510	30-day all cause readmission	3,512 features - demographics, admissions, diagnosis, procedure, medication, and lab data; unstructured text data from clinical notes	Deep unified neural networks, bag-of-words model for text	AUC 0.705 vs. 0.664 for LR	10-fold cross validation
Adler et al. (2019) ¹⁴⁴	Patients with initial occurrence of HF – N = 5,882	Death within 90 days (high-risk) or no HF event after 800 days (low-risk)	8 features - DBP, Cr, BUN, Hgb, WBC, Platelets, Albumin, RDW	GBM	AUC 0.88 (0.81 and 0.84 in external test cohorts)	50% random hold out validation, external test sets in in 2 separate cohorts
Angraal et al. (2019) ¹⁴⁵	Outpatients with HFpEF in TOPCAT – N = 1,767	All-cause mortality and HF Hosp through 3 years follow-up	86 features – demographic, clinical, laboratory, and electrocardiography data; KCCQ Scores	LR with LASSO, RF, GBM, SVM	RF (best) – AUC 0.72 mortality and 0.76 HF Hosp (vs. 0.66 and 0.73 for LRR)	5-fold cross-validation
Desai et al. (2020) ¹⁴⁶	Patients with HF by Medicare claims – N = 9,502	Mortality, HF Hosp, High Cost, Home Days Lost	54 features - demographics, HF-related variables, meds, comorbidities, and claims-based frailty-index and SES-index	LR with LASSO, CART, RF, GBM	AUC for GBM (best) vs. LR Mortality 0.727 vs. 0.7124 HF Hosp 0.745 vs. 0.707 High Cost 0.733 vs. 0.734 Home Days Lost 0.790 vs. 0.781	10-fold cross validation, Hold-out test set 3,389 patients

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Hosp = hospitalization, SES = socioeconomic status, DBP = diastolic blood pressure, Cr = creatinine, BUN = blood urea nitrogen, Hgb = hemoglobin, WBC = white blood cells, RDW = red cell distribution width, KCCQ = Kansas City Cardiomyopathy Questionnaire, RF = random forest, LR = logistic regression, GBM = gradient boosting machine, LASSO = least absolute shrinkage and selection operator, CART = classification and regression tree, SVM = support vector machine