



MACHINE LEARNING ALGORITHMS FOR PREDICTING HOSPITAL READMISSION AND MORTALITY RATES IN PATIENTS WITH HEART FAILURE

Rizinde, T.¹, Ngaruye, I.², and Cahill, N. D.³

¹ *Department of Applied Statistics, College of Business and Economics, University of Rwanda, Kigali, Rwanda.*

² *Department of Applied Mathematics, College of Science and Technology, University of Rwanda, Kigali, Rwanda.*

³ *School of Mathematics and Statistics, Rochester Institute of Technology, Rochester, NY 14623, USA.*

¹*rizindeth@gmail.com*

²*ingaruye@gmail.com*

³*ndcsma@rit.edu*

ABSTRACT

Purpose: The potential of predictive analytics in enhancing resource allocation and patient care for Heart failure (HF) outcomes is significant. This review aims to highlight this potential by analyzing existing studies and identifying the main barriers and challenges to applicability in all settings.

Design/ Methodology/ Approach: A comprehensive search of related articles was meticulously conducted across electronic databases, including Google Scholar, Web of Science, and PubMed. Using precise search phrases and keywords, 1,835 scholarly articles published between 1 January 2017 and 14 May 2024 were retrieved. Only 23 articles that met the strict inclusion criteria were considered, ensuring the validity of the findings. A quantitative meta-analysis approach was utilised.

Research Limitation/Implication: This research offers insights into enhancing healthcare outcomes as we analyse the challenges and feasibility of applying ML algorithms to predict heart failure outcomes in low-income settings.

Findings: The challenges include scalability, ethical and legal issues, the choice of appropriate ML model, interpretability, data availability, and healthcare professional mistrust of these ML algorithms.

Practical implications: This study offers practical strategies to bridge the gap between clinical practice and predictive analytics in these regions. These strategies should inspire and motivate healthcare professionals, researchers, and policymakers to consider and implement them.

Social implication: This study provides insights that may improve HF outcomes and healthcare delivery.

Originality/Value: The review identifies current gaps in the research, such as the need for more robust validation studies, the challenge of model interpretability, and the necessity for models that can be easily integrated into clinical workflows.

Keywords: *Heart failure. hospital. machine learning. mortality. readmission*



INTRODUCTION

Heart failure (HF) is a primary health concern globally that has a significant burden on patients and healthcare systems. At present, in 2024, HF is among the primary causes of hospital readmission and death, affecting over 64.34 million individuals globally. Of worldwide cases, 29% are classified as mild HF, 20% as moderate HF and 51% as severe HF (World Heart Federation, 2024). Heart failure can significantly lower a person's quality of life; more than half of all HF patients die within five years of receiving the diagnosis, mainly due to a lack of timely treatment (World Health Organization, 2024).

The International Congestive HF study examined HF mortality rates in various geographical locations, monitoring more than 5,820 patients for one year. A significant portion of the patients (42%) was in a critical condition (stage C or D) (Mpanya, Celik, Klug, & Ntsinjana, 2021a). Furthermore, while the overall patient mortality rate was 16.5% within one year, it was only 7% in China and 9% in South America. In contrast, the mortality rate in developing countries of Africa was found to be much higher at 34%, followed by 23% in Low and Middle-Income Countries (LMICs) of India (Dokainish et al., 2017). On top of that, the global increasing demand for healthcare due to increasing population, epidemics and new emerging diseases coupled with a shortage of medical professionals, high cost of healthcare, and unequal access to quality healthcare were also identified as significant challenges. This background highlights the necessity for better ways to identify high-risk patients earlier, for timely and targeted interventions to slow HF progression and improve the quality of life in all settings.

Prediction of HF patients' hospital readmission and mortality rates may enable healthcare professionals to identify high-risk patients, carry out prompt interventions, and improve patient outcomes. (Croon et al., 2022). These predictions are still challenging (Alabdjalbar et al., 2023). Historically, diagnosis, prognosis and other HF outcome predictions relied on traditional statistical methods supplemented by medical professionals' intuition, skills and experiences (Nduka et al., 2019) However, these traditional approaches often lead to unintentional biases and errors, and they are time-consuming and expensive, which may compromise the quality of services provided to patients (Uddin, Khan, Hossain, & Moni, 2019).

Despite the difficulties experienced in appropriately identifying HF outcomes, medical professionals rarely employ advanced predictive models in managing HF patients, especially in Low and Middle-Income Countries (LMICs) (Navarro et al., 2021). Nevertheless, the potential of machine learning (ML) to learn and recognise hidden patterns from large and complex data sets may enable the development of reliable prediction algorithms for HF hospital readmission and mortality rates (Golas et al., 2018; Hyland et al., 2020; Song et al., 2023). The benefits from life-saving technologies are not equally distributed to LMICs compared to the Western World (Savarese et al., 2023). Many of the challenges encountered in LMICs in the management of HF outcomes are related to human resource capacity, data availability and quality, and resource scarcity, especially in the form of financial limitations (Agbor et al., 2020a; Alabdjalbar et al.,



2023; Leong et al., 2023). Successful stories from developed countries affirm that the appropriate use of ML models for HF management has the potential to improve HF patients' outcomes (Adler et al., 2020a; Cho et al., 2021; Leong et al., 2023; Mpanya et al., 2021b). However, the existing literature lacks original research articles on the use of ML algorithms to manage HF outcomes that come from LMICs, particularly in Africa (Shahim, 2023). Therefore, it is essential and timely that a systematic approach be taken to understand the fundamental problems LMICs face.

This systematic literature review investigates the application of ML models to predict hospital readmission and mortality rates among HF patients from High, Middle, and Low-Income Countries. Furthermore, it examines existing studies to identify the main barriers and challenges that must be overcome to reduce the disparity in the advantages of these life-saving technologies for developed and developing countries.

UNIVERSAL DEFINITION, CLASSIFICATION AND GLOBAL BURDEN OF HEART FAILURE

Heart failure (HF) is traditionally defined as a condition characterised by reduced cardiac pumping capacity or impaired blood filling. (Shahim, 2023) An alternative definition is inadequate cardiac output brought on by abnormalities in structure or function or adequate cardiac output brought on by increased left ventricular filling pressure and compensatory neurohormonal activation. Although there are many definitions of HF, left ventricular ejection fraction (LVEF) has long been considered essential for HF diagnosis, characterisation, prognosis, and treatment selection. (Khan, Shahid, Fonarow, & Greene, 2022).

A universal and comprehensive definition and classification of heart failure (HF) was introduced in 2021. According to that definition, HF is a clinical syndrome characterised by signs or symptoms that derive from defects in the structure or function of the heart, boosted by high natriuretic peptide levels or objective evidence of pulmonary or systemic congestion. In addition, a revised HF staging includes four levels of people at risk for HF (Bozkurt et al., 2021). Stage A involves people who are at risk for HF but do not currently show any symptoms or indicators of HF. Based on structural or biomarker evaluations, these people do not show any signs of heart disease. In contrast, the second level or pre-HF (stage B), includes people who do not currently have symptoms or indicators of HF and have not had them in the past. These people show abnormal cardiac function, structural heart disease, or increased natriuretic peptide levels. On the other side, patients in stage C are people with current or prior HF symptoms or signs because of functional or structural heart abnormality. Lastly, patients with advanced HF (stage D) show severe symptoms and or signs of HF even at rest. These patients require advanced interventions such as transplant evaluation and mechanical circulatory support, or else they require palliative care, and they show resistance or intolerance to guideline-directed management and therapy (Bozkurt et al., 2021).

Similarly, HF has been reclassified into four types according to Left Ventricular Ejection Fraction (LVEF) ranges. These types include HF with preserved ejection fraction (HFpEF), or symptomatic



HF with left ventricular ejection fraction (LVEF) $\geq 50\%$; HF with mildly reduced ejection fraction (HFmrEF), or symptomatic HF with LVEF in between 41-49%; HF with reduced ejection fraction (HFrEF), or symptomatic HF with LVEF $\leq 40\%$; and HF with improved ejection fraction (HFmpEF), or symptomatic HF with a baseline LVEF of $\leq 40\%$, an increase of ≥ 10 points from baseline LVEF, and a second measurement of LVEF $> 40\%$ (Bozkurt et al., 2021).

With 90 % of the global burden of cardiovascular disease concentrated in LMICs, more than half of the global cases of HF are estimated to be shouldered by these countries in 2030 (Alabdjalbar et al., 2023). Unexpectedly, these countries have only 10 % of the world's research and healthcare resources to address this issue (Nakayama et al., 2023). As highlighted in the study by Shahim, (2023), despite a slight progress in HF prognosis in recent decades, the disease progresses to advanced stages for many HF patients, mainly in LMICs. In addition, the same research revealed a significant gap concerning HF epidemiology in developing countries outside Europe and North America. According to the same study, little available literature on these LMICs suggests a rapid rise in HF prevalence and associated consequences. It is a concern because the same study shows that younger people from low-income countries had a higher HF preference than in developed countries. Therefore, understanding the unique attributes of HF in these regions is paramount to developing appropriate machine-learning algorithms for predicting and classifying various HF outcomes.

Machine learning algorithms and statistical methods to predict HF outcomes

Prediction models identify the individuals most likely to be affected by a specific HF outcome and illuminate complex biological processes involved in the development and progression of HF under specific features. Research interest in modelling HF outcomes, such as rehospitalisation and mortality rates, has increased due to the abundance of available clinical data (Miao et al., 2018; Shameer et al., 2017). The existing literature has identified several methods being applied to predict HF outcomes in both LMICs and HICs. Several studies indicate that prediction models for different diseases are developed using machine learning and traditional statistical methods, although there is much overlap between the two (Cho et al., 2021; Sun et al., 2022). Traditional approaches emphasise examining a sample group to generalise about the population at large. In addition, standard statistics depend on programmers explicitly instructing the computer's rules (Sun et al., 2022). While these traditional rules-based approaches have the advantage of being transparent and easily comprehensible, the healthcare industry produces a vast amount of data that traditional methods cannot handle (Aldahiri et al., 2021). It becomes challenging to capture all critical information in the predetermined set of rules using these methods (Rajkomar et al., 2019). Moreover, these statistical methods have trouble with complex and nonlinear relationships found in most available data sets (Austin et al., 2022; Desai et al., 2020; Sun et al., 2022). These limitations substantially reduce effectiveness in complex medical scenarios. On the other hand, machine learning excels in learning from data samples (Alajmani & Jambi, 2020; Angraal et al., 2020a). As more data are fed into an ML model, it becomes more and more effective at predicting outcomes, eventually learning how to handle an issue without requiring all of the steps to be pre-



coded (Alajmani & Jambi, 2020). Because of this, ML is especially effective for tasks with complex and massive datasets or poorly understood relationships between variables.

After training an ML model, predictions can be made from new and unseen data points can be made (Beam & Kohane, 2018). Therefore, the decision to use statistical or ML models may be influenced by the available data, the project's overall objectives, and the model's intended level of interpretability. Today, in 2024, the health industry is producing enormous amounts of data, and ML is a substantial risk stratification technique for various disease outcome predictions.

Existing work on heart failure prediction models using machine-learning algorithms

Machine learning (ML) is a field within Artificial Intelligence (AI) that focuses on training computers to identify patterns from large datasets without being explicitly programmed. (Fregoso-Aparicio, Noguez, Montesinos, & García-García, 2021). Within this domain of ML, representation learning aims to find an accurate representation of knowledge obtained from data automatically. When this representation involves multiple layers of an artificial neural network, this is a deep learning. In shallow learning, data is usually transformed into a single layer of features. Deep learning expands on this idea by developing a hierarchy of features. Each layer contains patterns from the input, which exhibit greater abstraction as the layers progress from input to output.

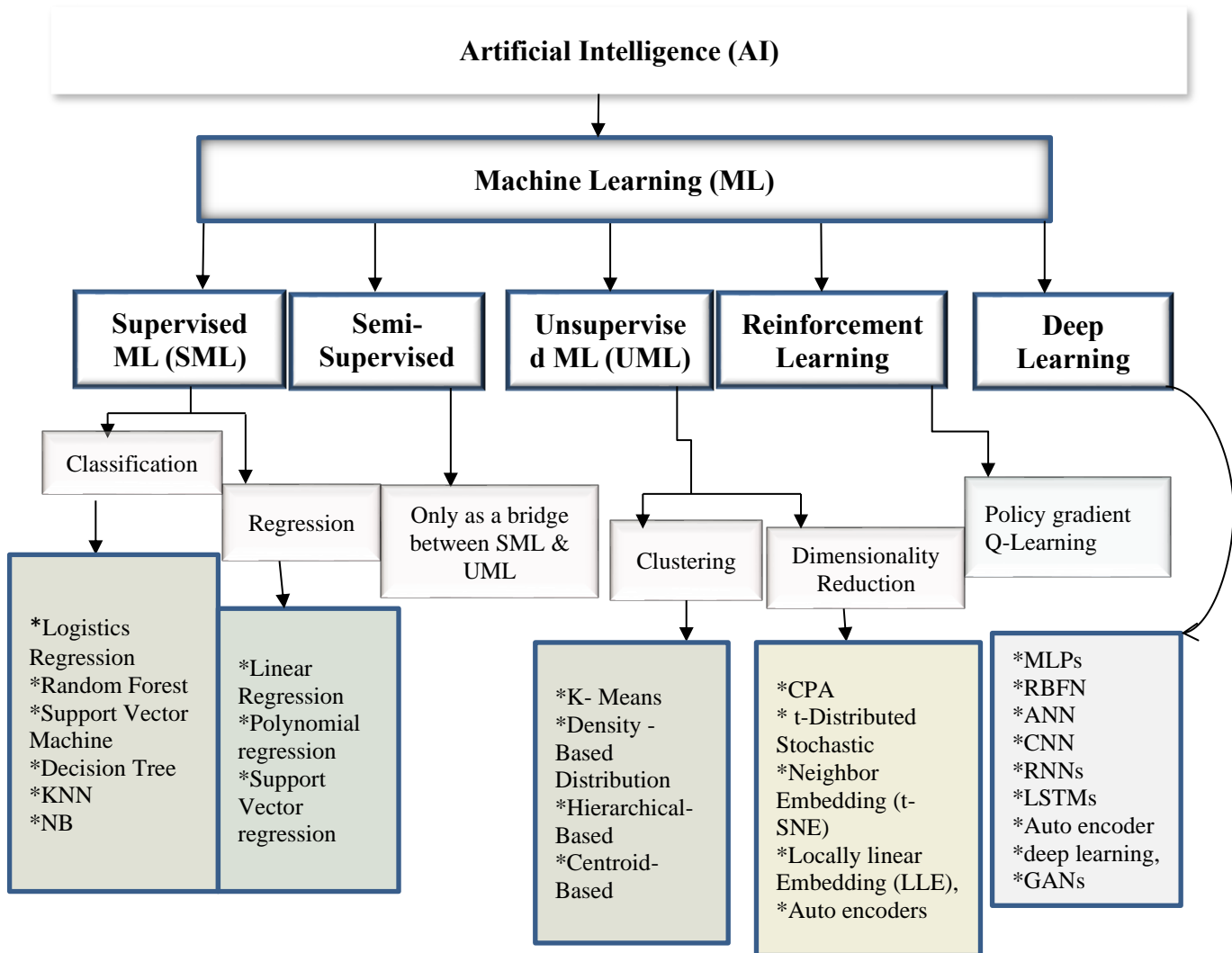
Several studies have been conducted on the application of ML models to predict different HF outcomes (Chen et al., 2019; Chicco & Jurman, 2020; Cho et al., 2021; Jagannatha & Yu, 2016b; Nakajima, Nakata, Doi, Tada, & Maruyama, 2020; Peirlinck et al., 2019; Zhao, Wood, Mirin, Cook, & Chunara, 2021). It has been shown that ML algorithms can learn complex hidden patterns in large datasets and that these algorithms can effectively predict HF outcomes such as hospital readmission and mortality rates (Saqib Ejaz Awan, Bennamoun, Sohel, Sanfilippo, & Dwivedi, 2019). For example, ML algorithms are used to generate risk scores for heart failure (Angraal et al., 2020a; Boodhun & Jayabalan, 2018; Chicco & Jurman, 2020; Kwon, Kim, Jeon, Lee, Lee, Cho, Choi, Jeon, Kim, & Kim, 2019; Yap, 2020). These risk scores predict the chance of an HF diagnosis and the probability of particular outcomes such as cause of mortality, cardiac death, hospitalisation and re-hospitalization, among others (Gheorghide et al., 2012).

In their study, Çolak et al. (2023), created a dataset of patient information containing 141 diseases with more than 2000 unique cases and 358 symptoms. Their paper used various ML models to predict the outcomes of different diseases, including HF. The authors witnessed that Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Naïve Bayes (NB), and K-Nearest Neighbor (KNN) are commonly utilised while predicting disease outcomes. Furthermore, they also examined Multilayer Perceptron (MLP), Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) algorithms. This is among the few studies that achieved the highest accuracy of 99.33% for a dataset of this size.



Again, Gandhi et al. (2020) using the dataset available at the portal of the Department of Biometrical Information (DBMI) of Columbia ('Columbia DBMI Home - Columbia DBMI', 2019), containing 42 types of diseases and 133 symptoms. They experimented with both supervised and unsupervised machine learning algorithms. Among these, they tested Logistic Regression (LR), Support Vector Machines (SVM), Random Forests (RF), Naïve Bayes (NB), Classification and Regression Trees (CART), K-Nearest Neighbors (KNN), and Linear Discriminant Analysis (LDA). The LR algorithm produced the lowest accuracy score, achieving 80.85%. Kumar et al. (Kumar, Sharma, & Prakash, 2021) Developed a Django application using ML algorithms to predict and avail clinical advice on general diseases, heart diseases and others. In general disease prediction, RF exhibited the highest accuracy score of 90.2% compared to the tested NB, LR, and KNN algorithms. On the other hand, LR was the most accurate at 92.3 % for heart disease prediction.

Fatima and Pasha (2017) investigated the diagnosis of heart disease using NB. They employed a dataset consisting of 500 patients from a diabetic institution in Chennai. Their results suggested that NB, despite its independence assumptions, might be a promising algorithm for heart disease prediction, offering an accuracy exceeding 86%. Furthermore, using structured heart disease patient and textbook data, Dhabarde et al., (2022). Dhabarde et al. (2022) created a dataset and used SVM, RF, NB, DT and LR algorithms for disease prediction. In their study, DT outperformed all other algorithms with an accuracy of 93.24%. Alanazi (2022) used the Convolutional Neural Networks (CNNs) and KNN algorithms for disease prediction on structured and unstructured real-world data. This model yielded a high accuracy of 95%, outperforming NB, LR, and DT in this study.



Adapted from the work of (Benti, Chaka, & Semie, 2023)
Figure 1: Main categories of machine learning algorithms

METHODOLOGY

This systematic review paper investigates different ML techniques in predicting HF outcomes, including hospital readmission and mortality rates. This study also assesses the effectiveness of ML models and explores their potential benefits and challenges while predicting these HF outcomes. The updated Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) standards are followed (Page et al., 2021). In addition, the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) serves as a reference for selecting relevant articles. A thorough and compressive search for articles published



between 1 January 2017 and 14 May 2024 was conducted across electronic databases, including MEDLINE, Google Scholar, EPUB, Embase, Web of Science, Scopus, Springer Link, and PubMed. Then, the search results were narrowed down to only articles relevant to this research by using the search phrases and keywords related to this study, such as HF prediction, ML, Artificial Intelligence (AI), Random Forests, Decision Trees, HF hospital readmission and mortality. The inclusive criteria involve original research published in peer-reviewed journals in the period mentioned earlier and in English. This review paper exclusively encompasses studies utilising ML models for prediction and classification, focusing on hospital readmission and mortality rates for patients diagnosed and hospitalised with HF from developed and developing countries. On the other hand, the exclusion criteria comprise all studies that do not focus on ML and HF patients and, therefore, do not report the outcomes of this research interest. In addition, all studies, including symposium or conference abstracts and articles published in predatory journals, were excluded. For any disagreement on selection, a team of three researchers reached a consensus.

Table1: Literature search keywords

Category	Search phrases	Keywords
Type of the study	Systematic review	Review, meta-analysis, systematic review,
Population	Patients hospitalised with heart failure.	Heart failure, congestive heart failure, decompensated heart failure, cardiac disease,
Intervention	Machine learning techniques	Artificial intelligence, Machine learning, Logistic regression, deep learning, Random forests, decision trees, support vector machine, Gradient, Artificial Neural Networks,
Outcomes	Hospital readmission, Mortality rate	Prehospitalisation, Readmission, Hospital stay,
	Mortality rates	Death, cardiovascular mortality, all-cause mortality,

Note that, to refine the search for this review, Boolean operators such as OR, AND, and NOT were used. For instance, Heart failure AND Random Forests AND hospital readmission. In addition, synonyms and related terms were also utilised, as far as appropriate indexing terms depending on the electronic databases used.



The review process

Initial results from the search produced 1,835 scholarly articles. Nevertheless, after rejecting 208 duplicates and excluding 1,367 due to their abstracts and titles, which were irrelevant to this research paper, only 260 scholarly articles remained. Most of these excluded works were theoretical studies and articles presented at conferences and symposia, and many did not focus specifically on HF and ML with predicted outcomes such as hospital readmission and mortality rates. After evaluating the eligibility of the remaining 260 full-text publications, 237 were disqualified due to many reasons, such as Outpatient Department (OPD), who were not admitted as inpatients, articles including patients with additional co-morbidities, which are not relevant to our study articles including only ischemic type of heart failure, and articles published in conferences. Twenty-three papers remained for analysis.

A modified version of the CHARMS checklist was developed and utilised to assess the quality of research and determine the likelihood of bias. The data that were abstracted for this study include, among other things, the study location and period, the year of publication and name of the author, the number of patients involved in the study, the source of data, the primary outcome, and the machine learning utilised. A quantitative meta-analysis approach was utilised.

FINDINGS AND DISCUSSION

Characteristics of the included studies

This review shows that almost 22% of included studies utilised data from only Electronic Health Records (EHR), and the other 35% used EHR but in combination with any other data such as administrative claims datasets or imaging data. Similarly, 9% of the analysed studies used only data collected and stored in patient registries, while 17% utilised registry data but in combination with any other type of dataset. Moreover, 9% of the studies used only trial data, while only 4% of the included studies utilised trial data in combination with another dataset. The remaining investigators involved the national inpatient sample and the MIMIC dataset.

Predictive models included in this research utilised sample sizes between 71 and 716,790, which is almost 6,701 patients on average per study. 26% of the studies predicted both death and hospital readmission of patients diagnosed with HF; similarly, 26 % predicted HF hospital readmission, whereas almost half (48%) of the included studies predicted HF mortality rates. Table 2 displays the period during which the data was obtained to create these prediction models.



Table 2: The characteristics of the included articles

Study ID	Data collection period	No. of Patients	Data Source	No. of features	Primary outcome	Machine learning algorithm (s)
Xu et al. (2024)	2018-2022	3,214	EHR	42	HF mortality rates	XG Boost, Random Forest
Angraal et al.(2020) (Angraal et al., 2020b)	2006-2013	1,767	Trial	26	HF re-hospitalization and mortality rates	Decision Trees
Desai et al. (2020)	2007-2014	9,502	Claims data and EHR	62	HF hospital readmission and mortality rates	Gradient Boosting
Wang et al. (2024)	2017-2020	8,921	Registry & EHR	29	All-cause HF mortality	Deep Neural Network (DNN)
Huang et al. (2021)	2014-2018	10,782	EHR	27	Early prediction of in-hospital HF mortality	Deep Learning Ensemble
Tian et al. (2023)	2019-2022	424	Registry data	67	Risk stratification for heart failure hospitalisation	Support Vector Machine (SVM), XG Boost
Singh et al. (2023)	2019-2022	5,289	EHR & Imaging data	102	HF hospitalisation	Convolutional Neural Network (CNN), Random Forest
Li et al. (2023)	2019-2021	1,540	Claims data & EHR	78	Hospital readmission (30-day)	Gradient Boosting, Support Vector Machine (SVM)
Zhou et al. (2021)	2015-2019	9,342	EHR	48	Hospital readmission or mortality	Support Vector Machine (SVM), Random Forest
Awan et al. (2019)	2003-2008	10,757	EHR	47	30 days HF hospital readmission and mortality rates,	Artificial Neural Network (ANN)
Gleeson et al. (2017)	2010-2015	295	Echo database and EHR	291	HF hospital readmissions	Random Forest



Hearn et al. (2018)	2001-2017	1,156	EHR and Cardiopulmonary	82	and mortality rates, HF mortality rates,	Logistic Regression
Khan et al. (2024)	2020-2023	2,417	Claims data	61	HF hospitalization or mortality	Support Vector Machine Logistic Regression
Jiang et al. (2019)	2013-2015	537	EHR	57	30-day hospital readmissions, HF	Random Forest
Lee et al. (2022) (Adler et al., 2020b)	2018-2021	3,145	Registry & Claims data	71	Long-term mortality	Cox Proportional Hazards Model, Support Vector Machine
Adler et al. (2020) (Adler et al., 2020a)	2006-2017	5,822	EHR/Trial	8	HF mortality	Logistic regression
Allam et al. (2019) (Allam et al., 2019)	2013	272, 778	Administrative Claim dataset	50	30 days readmission	Random forests
Afshaq et al.(2019) (2012-2016	7,655	EHR	36	30 days HF readmission	Naïve Bayes
Li et al. (2021)	2001 2012	1177	MIMIC-III database	20	HF In-hospital mortality	XG-Boost, LASSO regression
Chirinos et al. (2020)	2006-2012	379	Trial	48	Risk of all-cause death or heart failure-related hospital readmission	Logistic regression
Frizzell et al. (2017)	2005-2011	56,477	Registry and administrative claims Data	250	30 days HF readmission	Tree-augmented naïve Bayesian network, random forest, gradient-boosted, logistic regression.



Dai et al. (2017) (Dai et al., 2022)	2016-2019	4659	National inpatient Sample	34	HF in-hospital mortality	Random forest, Logistic regression
Kwon et al. (2019)	2016-2017	2165	Registry	32	12 and 36-month in-hospital HF mortality	Random forest, logistic regression, Bayesian network, and support vector machine,

Source: research work, 2024

Machine learning algorithms utilised in the included studies.

Just Twelve of the considered articles (52%) used a single method for building a predictive model, of which more than 25% utilised random forests, 8% utilised decision trees, 25% utilised Logistic Regression, 8% Naives Bayes, 8% Gradient boosting, 8% Deep Neural Network (DNN), 8% Deep Learning Ensemble. Conversely, 48% utilised at least two predictive models, including XG boost with Support Vector Machine and Decision Tree.

Predictors of the studies included in this review

In several studies included in this review, the authors stated only the total number of predictors (Morrill et al., 2022). Age, gender, diastolic blood pressure, left ventricular ejection fraction (LVEF), estimated glomerular filtration rate, haemoglobin, serum sodium, and blood urea nitrogen were all shown to be significant predictors of heart failure mortality. (Adler et al., 2020b; Golas et al., 2018). Nevertheless, ischemic cardiomyopathy, age, left ventricular ejection fraction, hypotension, haemoglobin, creatinine, and potassium serum levels were all shown to be predictors of hospital readmission (Desai et al., 2020).

Model development, internal and external validation

While developing a prediction model in machine learning, it is essential to mention that depending on the research project, there are differences in the methods of managing under-represented groups, data splitting, and internal and external validation details. According to the studies in this review, their initial data sets were split into two subsets. These include a training set between 60 and 80% of the original data and the testing set and internal validation set between 20 and 40 % of the initial data set (Ishaq et al., 2021). Although there was a dearth of data, only two of the thirty studies noted the need to validate the model externally. In addition, to handle the under-represented groups for training and testing, 83%. Twenty-five studies included in this review used Synthetic Minority Oversampling Techniques (SMOTE), 7% (2 studies) utilised importance weighting, and 10 % (3 studies) used stratified sampling. The SMOTE method takes the k-nearest-neighbour



strategy proposed by Chawla, Bowyer, Hall, and Kegelmeyer (2002), and randomly selects a neighbour among instances from under-represented classes. Then, instead of duplicating entries, new data points are generated based on pre-existing data points in the minority class to increase its representation (Xu et al., 2022).

Model performance and evaluation metrics

Based on the studies included in this systematic review, several classification algorithms have been employed to predict heart failure outcomes. These include Random Forests (RF), Logistic Regression (LR), Support vector machine (SVM), XGBoost, Naïve Bayes (NB) and Artificial Neural Networks (ANN) among others. (Alotaibi, 2019; Guidi et al., 2014; Guo et al., 2020; Jasinska-Piadlo et al., 2022; Shameer et al., 2017). The existing literature affirms that these methods efficiently estimate probabilities or make binary predictions by utilising features from the medical history and demographics of HF patients. On the other side, this review also involved studies that utilised survival analysis methods like Random Survival Forests and the Cox Proportional Hazards Model for predicting heart failure mortality risks over time and offering insights into patient prognosis. Furthermore, complementary techniques that address various data properties and problem kinds, such as Gradient Boosting and K-Nearest Neighbors, were also considered in this review (Golas et al., 2018; Stampehl et al., 2020).

The findings revealed that the evaluation metrics of the models were demonstrated in terms of their sensitivity, specificity, positive and negative predictive value, accuracy, and precision with the use of the confusion matrix (Kwon, et al., 2019; Mortazavi et al., 2016; Peirlinck et al., 2019; Plati et al., 2022). A confusion matrix, classification report, and other performance indicators have been extensively used to assess each model's efficacy. The classification report includes precision, recall, and f1-score for each class, while the confusion matrix displays the model's actual and predicted labels. Overall model accuracy, defined as the fraction of correct predictions, has also been reported in many studies involved in this review (Kwon, et al., 2019; Sun et al., 2022; Uddin et al., 2019). Several researchers also made use of the f-score, the Area Under the Receiver Operating Characteristics Curve (AUC), concordance statistic (C-statistic), and recall (Allam, Nagy, Thoma, & Krauthammer, 2019; Artetxe, Beristain, & Grana, 2018; Chicco & Jurman, 2020; Lorenzoni et al., 2019; Shin et al., 2021). The investigated prediction models for mortality yielded AUCs between 0.477 and 0.917, while those for hospital readmission had AUCs between 0.469 and 0.836 (Nakajima et al., 2020).

Numerous kinds of research examined in this review tried to evaluate how well various ML algorithms perform. However, it is challenging to determine which model can yield the best AUC, f1-score, or other metrics for any given study. This is because performance may depend on different factors, such as the data source, outcome predicted (hospital readmission, mortality rates, or both), candidate predictors, sample size and missing data, attrition, model construction, and model performance and evaluation.

In this review, it was found that the leading causes of unfavourable outcomes associated with HF and its treatment in LMICs are delayed diagnosis and presentation, co-existing conditions like



infections and malnutrition, a disorganised referral system, limited access to advanced diagnostic equipment, inadequately trained medical personnel, and a lack of financial support mechanisms among others. Again, it is highlighted that the application of ML models in these countries is facing systemic challenges, such as inadequate health informatics infrastructures and individual challenges, like restricted internet access and a shortage of smart devices and practical skills.

Discussion

As indicated by existing research, ischemic heart disease (IHD) was proven to be the primary cause of heart failure in high-income countries (HICs). However, hypertension remains the top cause in Low and Middle-Income Countries (LMICs), especially in sub-Saharan Africa (Agbor et al., 2020b; Bloomfield et al., 2013). In addition, as reported by several researchers and highlighted by the World Health Organization (WHO, 2024), there are significant gaps in the quality and accessibility of medical treatment across nations. Hence, risk calculators should be retrained using local data before being used in LMICs since algorithms trained on data from HICs may not be immediately relevant to LMICs. Not only is the New York Heart Association functional class only accessible in free-text format inside EHR systems, but machine learning methods used to produce risk ratings may not be entirely relevant in clinical settings owing to the lack of this crucial prognostic information. Nevertheless, recent developments, such as bidirectional long short-term memory with a layer of conditional random fields, have solved this problem. When developing and implementing risk ratings across geographic locations, it is essential to account for differences in patient populations and healthcare delivery systems (Jagannatha & Yu, 2016a). Table 3 displays a subset of the reviewed publications for this research.

Despite the high incidence of heart failure in LMICs, the findings revealed that few or no risk scores are provided to doctors and patients there (Guo et al., 2020). This is because there are insufficient centralised databases, registries, or multicenter studies to aggregate relevant data (Mpanya et al., 2021b). Instead of extrapolating data from high-income country research, LMICs might construct models to predict outcomes if they had access to digitally formatted health data. This group is underrepresented in the training and test datasets utilised in this systematic review due to the lack of standardised health data in LMICs. The AUC was mentioned as a measure of success in several of the research used to draw these conclusions. Using both clinical and physiological imaging data, the random forest method scored the most significant AUC up to 0.92 in several studies like the one by (Nakajima et al., 2020). Remember that if your model's AUC is below 0.50, you may as well flip a coin with your predictions since it cannot tell the difference between the classes. (Sun et al., 2022). There are some reasons why machine learning algorithms have only shown moderate effectiveness in specific applications. Among them are class imbalance, a lack of continuing communication between doctors and data scientists, and a training dataset with many missing data or few predictors. The study of healthcare data presents a challenging learning environment since antagonistic classes tend to outweigh positive classes, leaving fewer positive observations and patterns for algorithms to learn from. In the case of mortality prediction, for instance, it is standard for the class including dead patients to be smaller than the class containing live patients. (Turgeman & May, 2016).



The f-measure, also known as the f-Score or f1 Score, is equal to one for models with perfect accuracy and recall (Saito & Rehmsmeier, 2015). The percentage of true positives that were recognised is the metric for sensitivity. According to the research analysed, the sensitivity of machine learning algorithms ranged from 7.2% to 91.9%. Turgeman and May (2016) have shown that the sensitivity increases significantly when combining many prediction models into a single model using an ensemble technique. Although several studies have shown the random forest method to be highly accurate, followed by Logistic Regression, other studies have illustrated that it is not always the best choice for building HF-predicting models (Austin et al., 2022; Desai et al., 2020; Krumholz et al., 2019; Lorenzoni et al., 2019; Uddin et al., 2019). The findings showed that the main benefit of the random forest method is that it uses numerous decision tree algorithms on random data samples to form an ensemble-based classifier. Decision trees can quickly and clearly explain why a patient has been identified as high-risk to inform future risk reduction measures better. When numerous decision tree methods predict random data samples, the ensemble determines which class receives the most votes. Again, the findings confirmed that Random forests are also effective in dealing with the often-encountered problem of missing data in big healthcare datasets. They may also rate the accuracy of several forecasters (Uddin et al., 2019). Patients included in the models had a median age of 72, with the youngest being 40. In sub-Saharan African environments, where heart failure patients are generally a decade younger, this may restrict the usefulness of current risk calculators (Bloomfield et al., 2013; Glezeva et al., 2015).

Due to heart failure's heterogeneous etiologies and clinical manifestations, predictive models that include genetic, clinical, and imaging data are required. Our findings support the idea that doctors caring for patients with heart failure should prioritise the development of structured electronic health record (EHR) systems and the comparison of mortality and hospitalisation rates between patients treated with and without risk ratings. Clinicians without access to electronic health record systems should thoroughly investigate the cohort used to generate risk ratings before incorporating them into patient treatment (Nakajima et al., 2020).

CONCLUSION

In conclusion, the widespread adoption of risk calculators generated by machine learning is hampered by factors such as the variety of causes of heart failure, the scarcity of organised health information, the lack of trust in machine learning methods by medical professionals, and the moderate precision of predictive models.

Despite this caveat, the provided studies suggest that the random forests, logistic model, decision tree, KNN, and SVM linear with all kernel functions perform satisfactorily on the provided binary classification problem, with accuracy values between 0.71 and 0.93 and precision, recall, and f1-scores between 0.75 and 0.96. There is not a significant model class disparity between the two groups. Nevertheless, the decision tree model needs help correctly predicting class 2, as seen by the model's higher misclassifications for that group. The available data shows these models perform well on the specified classification task. This study explores how ML can be implemented to predict hospital readmissions and mortality for heart failure patients. By integrating these ML



tools into existing healthcare workflows, the research proposes a practical approach that could significantly improve patient outcomes and healthcare delivery's social and economic impacts.

REFERENCES

- Adler, E. D., Voors, A. A., Klein, L., Macheret, F., Braun, O. O., Urey, M. A., ... & Yagil, A. (2020a). Improving risk prediction in heart failure using machine learning. *European Journal of Heart Failure*, 22(1), 139–147. <https://doi.org/10.1002/ejhf.1628>
- Adler, E. D., Voors, A. A., Klein, L., Macheret, F., Braun, O. O., Urey, M. A., ... & Yagil, A. (2020b). Improving risk prediction in heart failure using machine learning. *European Journal of Heart Failure*, 22(1), 139–147. <https://doi.org/10.1002/ejhf.1628>
- Agbor, V. N., Ntusi, N. A. B., & Noubiap, J. J. (2020a). An overview of heart failure in low- and middle-income countries. *Cardiovascular Diagnosis and Therapy*, 10(2), 244–251. <https://doi.org/10.21037/cdt.2019.08.03>
- Agbor, V. N., Ntusi, N. A. B., & Noubiap, J. J. (2020b). An overview of heart failure in low- and middle-income countries. *Cardiovascular Diagnosis and Therapy*, 10(2), 24451–24251. <https://doi.org/10.21037/cdt.2019.08.03>
- Alabdajbar, M. S., Hasan, B., Noseworthy, P. A., Maalouf, J. F., Ammash, N. M., & Hashmi, S. K. (2023). Machine Learning in Cardiology: A Potential Real-World Solution in Low- and Middle-Income Countries. *Journal of Multidisciplinary Healthcare*, 16, 285–295. <https://doi.org/10.2147/JMDH.S383810>
- Alajmani, S., & Jambi, K. (2020). Assessing Advanced Machine Learning Techniques for Predicting Hospital Readmission. *International Journal of Advanced Computer Science and Applications*, 11(2). <https://doi.org/10.14569/IJACSA.2020.0110249>
- Alanazi, R. (2022). Identification and prediction of chronic diseases using machine learning approach. *Journal of Healthcare Engineering*, 2022. Retrieved from <https://downloads.hindawi.com/archive/2022/2826127.pdf>
- Aldahiri, A., Alrashed, B., & Hussain, W. (2021). Trends in Using IoT with Machine Learning in Health Prediction System. *Forecasting*, 3(1), 181–206. <https://doi.org/10.3390/forecast3010012>
- Allam, A., Nagy, M., Thoma, G., & Krauthammer, M. (2019). Neural networks versus Logistic regression for 30 days all-cause readmission prediction. *Scientific Reports*, 9(1), 9277.
- Alotaibi, F. S. (2019). Implementation of machine learning model to predict heart failure disease. *International Journal of Advanced Computer Science and Applications*, 10(6).
- Angraal, S., Mortazavi, B. J., Gupta, A., Khera, R., Ahmad, T., Desai, N. R., ... & Krumholz, H. M. (2020a). Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. *JACC: Heart Failure*, 8(1), 12–21.
- Angraal, S., Mortazavi, B. J., Gupta, A., Khera, R., Ahmad, T., Desai, N. R., ... & Krumholz, H. M. (2020b). Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. *JACC: Heart Failure*, 8(1), 12–21.
- Artetxe, A., Beristain, A., & Grana, M. (2018). Predictive models for hospital readmission risk: A systematic review of methods. *Computer Methods and Programs in Biomedicine*, 164, 49–64.



- Ashfaq, A., Sant'Anna, A., Lingman, M., & Nowaczyk, S. (2019). Readmission prediction using deep learning on electronic health records. *Journal of Biomedical Informatics*, 97, 103256.
- Austin, D. E., Lee, D. S., Wang, C. X., Ma, S., Wang, X., Porter, J., & Wang, B. (2022). Comparison of machine learning and the regression-based EHMRG model for predicting early mortality in acute heart failure. *International Journal of Cardiology*, 365, 78–84.
- Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. M., Chow, B. J., & Dwivedi, G. (2019). Feature selection and transformation by machine learning reduce variable numbers and improve prediction for heart failure readmission or death. *PloS one*, 14(6), e0218760.
- Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. M., & Dwivedi, G. (2019). Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. *ESC heart failure*, 6(2), 428-435.
- Beam, A. L., & Kohane, I. S. (2018). Big Data and Machine Learning in Health Care. *JAMA*, 319(13), 1317–1318. <https://doi.org/10.1001/jama.2017.18391>
- Benti, N. E., Chaka, M. D., & Semie, A. G. (2023). Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects. *Sustainability*, 15(9), 7087. <https://doi.org/10.3390/su15097087>
- Bloomfield, G. S., Barasa, F. A., Doll, J. A., & Velazquez, E. J. (2013). Heart failure in sub-Saharan Africa. *Current Cardiology Reviews*, 9(2), 157–173.
- Boodhun, N., & Jayabalan, M. (2018). Risk prediction in life insurance industry using supervised learning algorithms. *Complex & Intelligent Systems*, 4(2), 145–154.
- Bozkurt, B., Coats, A. J., Tsutsui, H., Abdelhamid, M., Adamopoulos, S., Albert, N., ... & Zieroth, S. (2021). Universal definition and classification of heart failure: A Report of the Heart Failure Society of America, Heart Failure Association of the European Society of Cardiology, Japanese Heart Failure Society and Writing Committee of the universal definition of heart failure. *Journal of cardiac failure*, 27(4), 387-413.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Chen, R., Lu, A., Wang, J., Ma, X., Zhao, L., Wu, W.,... & Yu, Z. (2019). Using machine learning to predict one-year cardiovascular events in patients with severe dilated cardiomyopathy. *European Journal of Radiology*, 117, 178–183.
- Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Medical Informatics and Decision Making*, 20(1), 1–16.
- Chirinos, J. A., Orlenko, A., Zhao, L., Basso, M. D., Cvijic, M. E., Li, Z., ... & Cappola, T. P. (2020). Multiple Plasma Biomarkers for Risk Stratification in Patients With Heart Failure and Preserved Ejection Fraction. *Journal of the American College of Cardiology*, 75(11), 1281–1295. <https://doi.org/10.1016/j.jacc.2019.12.069>
- Cho, S. M., Austin, P. C., Ross, H. J., Abdel-Qadir, H., Chicco, D., Tomlinson, G., ... & Billia, F. (2021). Machine learning compared with conventional statistical models for predicting



- myocardial infarction readmission and mortality: A systematic review. *Canadian Journal of Cardiology*, 37(8), 1207–1214.
- Çolak, M., Sivri, T. T., Akman, N. P., Berkol, A., & Ekici, Y. (2023). Disease prognosis using machine learning algorithms based on new clinical dataset. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, 65(1), 52–68.
- Columbia DBMI Home—Columbia DBMI. (2019, June 11). Retrieved 10 April 2024, from <https://www.dbmi.columbia.edu/>, <https://www.dbmi.columbia.edu/>
- Croon, P. M., Selder, J. L., Allaart, C. P., Bleijendaal, H., Chamuleau, S. A. J., Hofstra, L., & Winter, M. M. (2022). Current state of artificial intelligence-based algorithms for hospital admission prediction in patients with heart failure: A scoping review. *European Heart Journal - Digital Health*, 3(3), 415–425. <https://doi.org/10.1093/ehjdh/ztac035>
- Dai, Q., Sherif, A. A., Jin, C., Chen, Y., Cai, P., & Li, P. (2022). Machine learning predicting mortality in sarcoidosis patients admitted for acute heart failure. *Cardiovascular Digital Health Journal*, 3(6), 297–304. <https://doi.org/10.1016/j.cvdhj.2022.08.001>
- Desai, R. J., Wang, S. V., Vaduganathan, M., Evers, T., & Schneeweiss, S. (2020). Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. *JAMA Network Open*, 3(1), e1918962–e1918962.
- Dhabarde, S., Mahajan, R., Mishra, S., Chaudhari, S., Manelu, S., & Shelke, N. S. (2022). Disease prediction using machine learning algorithms. *International Research Journal of Modernization in Engineering Technology and Science*, 4, 379–384.
- Dokainish, H., Teo, K., Zhu, J., Roy, A., AlHabib, K. F., ElSayed, A.,... & Mondo, C. (2017). Global mortality variations in patients with heart failure: Results from the International Congestive Heart Failure (INTER-CHF) prospective cohort study. *The Lancet Global Health*, 5(7), e665–e672. [https://doi.org/10.1016/S2214-109X\(17\)30196-1](https://doi.org/10.1016/S2214-109X(17)30196-1)
- Fatima, M., & Pasha, M. (2017). Survey of Machine Learning Algorithms for Disease Diagnostic. *Journal of Intelligent Learning Systems and Applications*, 9(1), 1–16. <https://doi.org/10.4236/jilsa.2017.91001>
- Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: A systematic review. *Diabetology & Metabolic Syndrome*, 13(1), 148. <https://doi.org/10.1186/s13098-021-00767-9>
- Frizzell, J. D., Liang, L., Schulte, P. J., Yancy, C. W., Heidenreich, P. A., Hernandez, A. F.,... & Laskey, W. K. (2017). Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: Comparison of machine learning and other statistical approaches. *JAMA Cardiology*, 2(2), 204–209.
- Gandhi, K., Mittal, M., Gupta, N., & Dhall, S. (2020). Disease prediction using machine learning. *International Journal for Research in Applied Science and Engineering Technology*, 8(6). Retrieved from <https://www.academia.edu/download/63686168/7720200620-126547-18e0xrm.pdf>



- Gheorghiadu, M., Albert, N. M., Curtis, A. B., Thomas Heywood, J., McBride, M. L., Inge, P. J., ... & Fonarow, G. C. (2012). Medication dosing in outpatients with heart failure after implementation of a practice-based performance improvement intervention: findings from IMPROVE HF. *Congestive Heart Failure*, 18(1), 9-17.
- Gleeson, S., Liao, Y.-W., Dugo, C., Cave, A., Zhou, L., Ayar, Z., ... & Gavin, A. (2017). ECG-derived spatial QRS-T angle is associated with ICD implantation, mortality and heart failure admissions in patients with LV systolic dysfunction. *PloS One*, 12(3), e0171069.
- Glezeva, N., Gallagher, J., Ledwidge, M., O'Donoghue, J., McDonald, K., Chipolombwe, J., & Watson, C. (2015). Heart failure in sub-Saharan Africa: Review of the aetiology of heart failure and the role of point-of-care biomarker diagnostics. *Tropical Medicine and International Health*, 20(5), 8.
- Golas, S. B., Shibahara, T., Agboola, S., Otaki, H., Sato, J., Nakae, T., ... & Jethwani, K. (2018). 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 Medical Informatics and Decision Making*, 18(1), 44. <https://doi.org/10.1186/s12911-018-0620-z>
- Guidi, G., Pettenati, M. C., Melillo, P., & Iadanza, E. (2014). A machine learning system to improve heart failure patient assistance. *IEEE Journal of Biomedical and Health Informatics*, 18(6), 1750–1756.
- Guo, A., Pasque, M., Loh, F., Mann, D. L., & Payne, P. R. (2020). Heart failure diagnosis, readmission, and mortality prediction using machine learning and artificial intelligence models. *Current Epidemiology Reports*, 7, 212–219.
- Hearn, J., Ross, H. J., Mueller, B., Fan, C.-P., Crowdy, E., Duhamel, J., ... & Manlhiot, C. (2018). Neural networks for prognostication of patients With heart failure: Improving performance through the incorporation of breath-by-breath data from cardiopulmonary exercise testing. *Circulation: Heart Failure*, 11(8), e005193.
- Hyland, S. L., Faltys, M., Hüser, M., Lyu, X., Gumbsch, T., Esteban, C., ... & Merz, T. M. (2020). Early prediction of circulatory failure in the intensive care unit using machine learning. *Nature Medicine*, 26(3), 364–373. <https://doi.org/10.1038/s41591-020-0789-4>
- Ishaq, A., Sadiq, S., Umer, M., Ullah, S., Mirjalili, S., Rupapara, V., & Nappi, M. (2021). Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques. *IEEE Access*, 9, 39707–39716. <https://doi.org/10.1109/ACCESS.2021.3064084>
- Jagannatha, A. N., & Yu, H. (2016a). Bidirectional RNN for medical event detection in electronic health records. *Proceedings of the Conference. Association for Computational Linguistics. North American Chapter. Meeting, 2016*, 473. NIH Public Access.
- Jagannatha, A. N., & Yu, H. (2016b). Structured prediction models for RNN based sequence labeling in clinical text. *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, 2016*, 856. NIH Public Access.
- Jasinska-Piadlo, A., Bond, R., Biglarbeigi, P., Brisk, R., Campbell, P., & McEneaney, D. (2022). What can machines learn about heart failure? A systematic literature review. *International Journal of Data Science and Analytics*, 13(3), 163–183.



- Jiang, W., Siddiqui, S., Barnes, S., Barouch, L. A., Korley, F., Martinez, D. A., ... & Levin, S. (2019). Readmission Risk Trajectories for Patients With Heart Failure Using a Dynamic Prediction Approach: Retrospective Study. *JMIR Medical Informatics*, 7(4), e14756. <https://doi.org/10.2196/14756>
- Khan, M. S., Shahid, I., Fonarow, G. C., & Greene, S. J. (2022). Classifying heart failure based on ejection fraction: Imperfect but enduring. *European Journal of Heart Failure*, 24(7), 1154–1157. <https://doi.org/10.1002/ejhf.2470>
- Krumholz, H. M., Coppi, A. C., Warner, F., Triche, E. W., Li, S.-X., Mahajan, S., ... & Dorsey, K. (2019). Comparative effectiveness of new approaches to improve mortality risk models from Medicare claims data. *JAMA Network Open*, 2(7), e197314–e197314.
- Kumar, A., Sharma, G. K., & Prakash, U. M. (2021). Disease prediction and doctor recommendation system using machine learning approaches. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 9, 34–44.
- Kwon, J., Kim, K.-H., Jeon, K.-H., Lee, S. E., Lee, H.-Y., Cho, H.-J., ... & Kim, J.-J. (2019). Artificial intelligence algorithm for predicting mortality of patients with acute heart failure. *PloS One*, 14(7), e0219302.
- Kwon, J. M., Kim, K. H., Jeon, K. H., Lee, S. E., Lee, H. Y., Cho, H. J., ... & Oh, B. H. (2019). Artificial intelligence algorithm for predicting mortality of patients with acute heart failure. *PloS one*, 14(7), e0219302.
- Leong, D. P., Joseph, P., McMurray, J. J. V., Rouleau, J., Maggioni, A. P., Lanan, F., ... & (2023). Frailty and outcomes in heart failure patients from high-, middle-, and low-income countries. *European Heart Journal*, 44(42), 4435–4444. <https://doi.org/10.1093/eurheartj/ehad595>
- Li, F., Xin, H., Zhang, J., Fu, M., Zhou, J., & Lian, Z. (2021). Prediction model of in-hospital mortality in intensive care unit patients with heart failure: Machine learning-based, retrospective analysis of the MIMIC-III database. *BMJ Open*, 11(7), e044779. <https://doi.org/10.1136/bmjopen-2020-044779>
- Li, J., Li, X., Liu, H., Gao, L., Wang, W., Wang, Z., ... & Wang, Q. (2023). Climate change impacts on wastewater infrastructure: A systematic review and typological adaptation strategy. *Water Research*, 242, 120282. <https://doi.org/10.1016/j.watres.2023.120282>
- Lorenzoni, G., Sabato, S. S., Lanera, C., Bottigliengo, D., Minto, C., Ocagli, H., ... & Pisanò, F. (2019). Comparison of machine learning techniques for prediction of hospitalization in heart failure patients. *Journal of Clinical Medicine*, 8(9), 1298.
- Miao, F., Cai, Y.-P., Zhang, Y.-X., Fan, X.-M., & Li, Y. (2018). Predictive modeling of hospital mortality for patients with heart failure by using an improved random survival forest. *IEEE Access*, 6, 7244–7253.
- Morrill, J., Qirko, K., Kelly, J., Ambrosy, A., Toro, B., Smith, T., ... & Swaminathan, S. (2022). A Machine Learning Methodology for Identification and Triage of Heart Failure Exacerbations. *Journal of Cardiovascular Translational Research*, 15(1), 103–115. <https://doi.org/10.1007/s12265-021-10151-7>



- Mortazavi, B. J., Downing, N. S., Bucholz, E. M., Dharmarajan, K., Manhapra, A., Li, S.-X., & Krumholz, H. M. (2016). Analysis of machine learning techniques for heart failure readmissions. *Circulation: Cardiovascular Quality and Outcomes*, 9(6), 629–640.
- Mpanya, D., Celik, T., Klug, E., & Ntsinjana, H. (2021a). Machine learning and statistical methods for predicting mortality in heart failure. *Heart Failure Reviews*, 26(3), 545–552.
- Mpanya, D., Celik, T., Klug, E., & Ntsinjana, H. (2021b). Predicting mortality and hospitalization in heart failure using machine learning: A systematic literature review. *IJC Heart & Vasculature*, 34, 100773.
- Nakajima, K., Nakata, T., Doi, T., Tada, H., & Maruyama, K. (2020). Machine learning-based risk model using 123 I-metaiodobenzylguanidine to differentially predict modes of cardiac death in heart failure. *Journal of Nuclear Cardiology*, 1–12.
- Nakayama, L. F., Zago Ribeiro, L., Novaes, F., Miyawaki, I. A., Miyawaki, A. E., de Oliveira, J. A. E., ... & Silva, P. S. (2023). Artificial intelligence for telemedicine diabetic retinopathy screening: a review. *Annals of Medicine*, 55(2), 2258149.
- Navarro, C. L. A., Damen, J. A. A., Takada, T., Nijman, S. W. J., Dhiman, P., Ma, J., ... & Hooft, L. (2021). Risk of bias in studies on prediction models developed using supervised machine learning techniques: Systematic review. *BMJ*, 375, n2281. <https://doi.org/10.1136/bmj.n2281>
- Nduka, A., Samual, J., Elango, S., Divakaran, S., Umar, U., & SenthilPrabha, R. (2019). Internet of Things Based Remote Health Monitoring System Using Arduino. *2019 Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 572–576. <https://doi.org/10.1109/I-SMAC47947.2019.9032438>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Peirlinck, M., Sahli Costabal, F., Sack, K. L., Choy, J. S., Kassab, G. S., Guccione, J. M., & Kuhl, E. (2019). Using machine learning to characterize heart failure across the scales. *Biomechanics and Modeling in Mechanobiology*, 18, 1987–2001.
- Plati, D. K., Tripoliti, E. E., Karanasiou, G. S., Rammos, A., Bechlioulis, A., Watson, C. J., & Naka, K. K. (2022). Machine learning techniques for predicting and managing heart failure. *Predicting Heart Failure: Invasive, Non-Invasive, Machine Learning and Artificial Intelligence Based Methods*, 189–226.
- Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347-1358.
- Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS One*, 10(3), e0118432.
- Savarese, G., Becher, P. M., Lund, L. H., Seferovic, P., Rosano, G. M. C., & Coats, A. J. S. (2023). Global burden of heart failure: A comprehensive and updated review of epidemiology. *Cardiovascular Research*, 118(17), 3272–3287. <https://doi.org/10.1093/cvr/cvac013>



- Shahim, B. (2023). *Global Public Health Burden of Heart Failure: An Updated Review*. Retrieved from <https://www.cfrjournal.com/articles/global-public-health-burden-heart-failure-updated-review>
- Shameer, K., Johnson, K. W., Yahi, A., Miotto, R., Li, L. I., Ricks, D., ... & Gelijns, S. (2017). Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: A case-study using Mount Sinai heart failure cohort. *Pacific Symposium on Biocomputing 2017*, 276–287. World Scientific.
- Shin, S., Austin, P. C., Ross, H. J., Abdel-Qadir, H., Freitas, C., Tomlinson, G., ... & Billia, F. (2021). Machine learning vs. Conventional statistical models for predicting heart failure readmission and mortality. *ESC Heart Failure*, 8(1), 106–115.
- Song, X., Tong, Y., Luo, Y., Chang, H., Gao, G., Dong, Z., ... & Tong, R. (2023). Predicting 7-day unplanned readmission in elderly patients with coronary heart disease using machine learning. *Frontiers in Cardiovascular Medicine*, 10. <https://doi.org/10.3389/fcvm.2023.1190038>
- Stampehl, M., Friedman, H. S., Navaratnam, P., Russo, P., Park, S., & Obi, E. N. (2020). Risk assessment of post-discharge mortality among recently hospitalized Medicare heart failure patients with reduced or preserved ejection fraction. *Current Medical Research and Opinion*, 36(2), 179–188.
- Sun, Z., Dong, W., Shi, H., Ma, H., Cheng, L., & Huang, Z. (2022). Comparing machine learning models and statistical models for predicting heart failure events: A systematic review and meta-analysis. *Frontiers in Cardiovascular Medicine*, 9, 647.
- Tian, P., Liang, L., Zhao, X., Huang, B., Feng, J., Huang, L., ... & Zhang, Y. (2023). Machine Learning for Mortality Prediction in Patients With Heart Failure With Mildly Reduced Ejection Fraction. *Journal of the American Heart Association*, 12(12), e029124. <https://doi.org/10.1161/JAHA.122.029124>
- Turgeman, L., & May, J. H. (2016). A mixed-ensemble model for hospital readmission. *Artificial Intelligence in Medicine*, 72, 72–82.
- Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, 19(1), 1–16.
- Wang, H., Li, Y., Chai, K., Long, Z., Yang, Z., Du, M., ... & Yang, J. (2024). Mortality in patients admitted to hospital with heart failure in China: A nationwide Cardiovascular Association Database-Heart Failure Centre Registry cohort study. *The Lancet Global Health*, 12(4), e611–e622. [https://doi.org/10.1016/S2214-109X\(23\)00605-8](https://doi.org/10.1016/S2214-109X(23)00605-8)
- WHO. (2024). Health equity. Retrieved 16 March 2024, from https://www.who.int/health-topics/health-equity#tab=tab_1
- World Health Organization. (2024). Cardiovascular diseases. Retrieved 13 March 2024, from <https://www.who.int/health-topics/cardiovascular-diseases>
- World Heart Federation. (2024). Heart Failure | What We Do. Retrieved 13 March 2024, from World Heart Federation website: <https://world-heart-federation.org/what-we-do/heart-failure/>



- Xu, L., Cao, F., Wang, L., Liu, W., Gao, M., Zhang, L., ... & Lin, M. (2024). Machine learning model and nomogram to predict the risk of heart failure hospitalization in peritoneal dialysis patients. *Renal Failure*, 46(1), 2324071.
<https://doi.org/10.1080/0886022X.2024.2324071>
- Xu, Z., Shen, D., Kou, Y., & Nie, T. (2022). A Synthetic Minority Oversampling Technique Based on Gaussian Mixture Model Filtering for Imbalanced Data Classification. *IEEE Transactions on Neural Networks and Learning Systems*, 1–14.
<https://doi.org/10.1109/TNNLS.2022.3197156>
- Yap, J. (2020). Risk stratification in heart failure: Existing challenges and potential promise. *International Journal of Cardiology*, 313, 97–98.
- Zhao, Y., Wood, E. P., Mirin, N., Cook, S. H., & Chunara, R. (2021). Social determinants in machine learning cardiovascular disease prediction models: A systematic review. *American Journal of Preventive Medicine*, 61(4), 596–605.