DEVELOPMENTAL MEDICO-LIFE-SCIENCES

ISSN (P): 3007-2786, (E): 3007-2794 DOI: https://doi.org/10.69750/dmls.03.06.0217

ORIGINAL RESEARCH ARTICLE Open Access

Machine Learning Versus Traditional Risk Scores for Predicting In-Hospital Mortality in Acute Heart Failure

Ameer Hamza[1], Sana Yousaf[1], Hamza Akram[2], Muhammad Kashif[2]

1 Sheikh Zayed Medical College, Rahim Yar Khan, Pakistan.
2 University of Hertfordshire, Hatfield, AL10 9AB, United Kingdom.

Corresponding Author: Hamza Akram, **Email:** chhamzajalal0300@gmail.com

ABSTRACT

Background: However, early prediction of in-hospital mortality in acute heart failure needs to be accurate for timely escalation of care. Traditional risk scores are helpful but are not necessarily able to detect the complex nonlinear interactions between clinical and laboratory variables.

Objective: To assess machine-learning algorithms in comparison to traditional risk assessments for predicting in-hospital mortality in patients admitted with acute heart failure.

Methods: This prospective observational study was conducted from January 2025 to August 2025, including 120 consecutive individuals identified with acute heart failure. Traditional risk evaluation was performed with ADHERE, GWTG-HF, and OPTIMIZE-HF metrics. This study used logistic regression, support vector machines, Random Forest, LightGBM, XGBoost, and artificial neural networks as machine learning algorithms. The efficacy of the models was evaluated by AUC, accuracy, sensitivity, specificity, predictive values, and F1-score.

Results: 18 patients (15.0%) died in hospital. Non-survivors were older and had lower blood pressure, oxygen saturation, higher creatinine, hyponatremia, elevated lactate, reduced left ventricular ejection fraction, cardiogenic shock, and greater inotropic requirement. GWTG-HF had the highest AUC value when compared to traditional scores (0.78). Among the three models, XGBoost outperformed the others with the largest AUC of 0.91, followed by LightGBM and Random Forest, with values of 88.3%, 83.3% and 89.2% respectively. Systolic blood pressure, serum creatinine, lactate, oxygen saturation, and left ventricular ejection fraction were the best predictors.

Conclusion: The machine-learning models, especially XGBoost, predicted in-hospital mortality in acute heart failure better than traditional risk scores.

Keywords: Acute heart failure; machine learning; XGBoost; mortality; risk prediction.

Received: 10/03/2026
Revised: 25/05/2026
Accepted: 26/06/2026
Published: 30/06/2026

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third-party material in this article are included in the article's Creative Commons license unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you must obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/public_domain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

INTRODUCTION

Acute heart failure (AHF) is a prevalent cardiovascular emergency and a primary cause of hospital admissions among people globally [1]. Nonetheless, despite substantial advancements in pharmacological therapy, interventional cardiology, and multidisciplinary management of heart failure, it continues to be a significant contributor to in-hospital mortality, prolonged length of stay, and healthcare costs [2]. Individuals with acute heart failure (AHF) have a wide spectrum of clinical symptoms, ranging from mild congestion to profound cardiogenic shock. It is essential to recognize patients who are at increased risk for negative outcomes in their treatment management. Precise forecasting of in-hospital mortality allows physicians to select patients for intensive observation, expedite necessary therapeutic interventions, enhance resource allocation, and provide suitable patient guidance.

The categorization of risk has gained significance in the treatment of acute heart failure. Numerous recognized clinical risk assessments, such as the Acute Decompensated Heart Failure National Registry (ADHERE), Get with The Guidelines–Heart Failure (GWTG-HF), and OPTIMIZE-HF scores, have been developed to predict short-term mortality based on fundamental clinical and laboratory information [4].

Such devices have proven to have acceptable predictive power and are widely adopted due to their ease and simplicity of use and their seamless integration into clinical care. Most of these models are based on age, selected comorbid conditions, serum sodium concentration, heart rate, renal function, and systolic blood pressure, which are used to estimate the probability of adverse clinical outcomes during hospitalization [5,6].

While traditional risk scores have proven to be clinically useful, they also have several important limitations. Most have been created with traditional statistical regression-based techniques that rely on linear relationships between predictor variables and clinical outcomes [7]. The pathophysiology of acute heart failure, however, is very complicated, with dynamic interactions between cardiac function, renal derangement, activation of various neurohormones, systemic inflammation, metabolic disturbances, and several co-existing illnesses. The relationships between these variables are complex and nonlinear, and may not be easily represented by traditional statistical models. In addition, traditional scoring systems tend to have few predictor variables and often lack the ability to adjust for temporal changes in physiological variables, biomarker profiles, treatment responses, or complex interactions between clinical variables, which may affect their predictive power in patient groups with diverse characteristics [8,9].

The fast-paced development of digital healthcare systems and electronic medical records has significantly propelled the adoption of AI in cardiovascular medicine. Of these technologies, machine learning has shown potential for clinical prediction, owing to its capacity to analyse large multidimensional data sets and to detect complex patterns that might not be discerned using conventional statistical methods. Machine-learning algorithms can process many variables at once, model nonlinear relations, discover hidden interactions, and continuously improve their predictive performance with iterative learning processes. Machine learning is especially well suited to predict outcomes in acute heart failure, where prognosis is multifactorial, depending on the interplay of demographic, clinical, laboratory, imaging, treatment, and physiological response [10,11].

There have been several machine learning algorithms that have shown promising performance in cardiovascular risk prediction. Random Forest, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), CatBoost, artificial neural networks, and support vector machines have been demonstrated to be better at discriminating in predicting mortality, intensive care unit (ICU) admission, rehospitalisation, and other adverse cardiovascular events. Such algorithms can capture vast amounts of clinical data and detect intricate relationships between the data that may not be captured by conventional prediction models. Beyond this, improvements in explainability of AI models have made it easier for clinicians to interpret what parts of a patient's data contribute to the estimation of risk using machine-learning models [12-15].

Although there is a rising interest in artificial intelligence, the use of machine learning models in the clinical setting of acute heart failure remains scarce. Retrospective data sets or administrative databases have been used in many published studies, or the patients included in the study may have been highly selected and do not fully represent standard patient care. There is also a large degree of variation in the model development, predictor selection, validation process, and definition of outcomes, rendering it difficult to make direct comparisons between studies. Furthermore, few studies have directly and prospectively performed comparisons of machine learning algorithms to well-known traditional risk scores in the same patient population with a uniform measurement tool [16,17].

In resource-constrained health care environments, better prognostic tools are especially timely, as early identification of high-risk patients could be useful for more efficient use of intensive care (ICU) resources and for clinical decision-making. Prediction models based on routinely collected clinical data plus sophisticated computational techniques might provide improved personalised care for patients and evidence-based management. Before their large-scale use in clinical workflows, these approaches need to be tested in prospective validation studies [18,19].

Thus, the current study aimed to compare the predictive capacity of machine-learning algorithms with traditional clinical risk scores for in-hospital mortality in acute heart failure patients [20]. The current study aims to compare the discrimination, calibration, sensitivity, specificity, overall predictive accuracy, and clinical utility of both approaches based on a prospectively collected clinical and laboratory dataset. The hypothesis was that machine-learning models would be better at predicting outcomes and at-risk stratification than conventional risk scoring systems and could therefore be useful in improving early clinical decision-making and optimizing patient outcomes in acute heart failure.

MATERIALS AND METHODS

This prospective observational predictive study was carried out at Sheikh Zayed Hospital in Rahim Yar Khan, Pakistan, from January 2025 to August 2025. Ethical Approval Reference No.: SZMC/IRB/2024/127 was

granted from the Institutional Review Board (IRB) of Sheikh Zayed Medical College, Rahim Yar Khan, for the research. All participants or their legally designated representatives furnished signed informed consent before to registration. The objective of the study was to evaluate the efficacy of machine-learning models in forecasting in-hospital mortality during the first hospitalization in comparison to conventional clinical risk assessments.

A non-probability sampling technique was used to enlist 120 adult individuals diagnosed with acute heart failure. Acute heart failure is defined as the onset or worsening of symptoms and signs of heart failure, requiring immediate hospitalization and intravenous therapy. Inclusion criteria were individuals over 18 years old who were hospitalized for acute heart failure (AHF) and had comprehensive baseline clinical and laboratory information during the first 24 hours of admission, receiving continuous hospital care until discharge or demise.

Patients with incomplete medical records, who left AMA, who were transferred to another hospital before outcome assessment, who had terminal non-cardiac illness (expected survival less than 24 hours), who required immediate intervention before complete baseline assessment, and who declined consent were excluded.

The data were obtained using a structured proforma, prospectively. The baseline demographic factors were age, sex, body mass index (BMI), smoking status, diabetes mellitus, hypertension, chronic renal disease, ischemic heart disease, atrial fibrillation, and previous admissions for heart failure. Clinical factors at presentation were systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, oxygen saturation, New York Heart Association functional class, peripheral edema, pulmonary crepitations, oxygen demand, and need for inotropic support.

Laboratory parameters obtained upon admission within the initial 24 hours included: hemoglobin concentration, total leukocyte count, serum sodium, serum potassium, serum creatinine, blood urea nitrogen, estimated glomerular filtration rate (GFR), random blood glucose, serum albumin, C-reactive protein, lactate, cardiac troponin, and B-type natriuretic peptide (BNP) or N-terminal pro-B-type natriuretic peptide (NT-proBNP) when accessible. The echocardiographic data included left ventricular ejection fraction, right ventricular dysfunction, severe valvular disease, and estimated pulmonary artery pressure.

Using accepted acute heart failure risk scores, traditional risk stratification was done for each patient on the ADHERE risk tree, GWTG-HF risk score, and OPTIMIZE-HF risk score. Admission data and laboratory tests at baseline were used to obtain the necessary information for calculating each risk score. To minimize observer bias, risk scores were determined before outcome analysis.

All clinical, laboratory, and echocardiographic parameters were considered for machine-learning model development. Data were cleaned before training the model. Variables with a large number of missing values were excluded; median imputation for continuous and mode imputation for categorical values were performed for the missing values. Where necessary, continuous variables were standardized based on the assumptions of the algorithm used.

The entire dataset was split into training and testing sets with 80% of patients in the training set and 20% in the testing set. Models were constructed and internally optimized using the training dataset, and the testing dataset was used for the final assessment of performance. Repeated cross-validation was also done to mitigate the risk of overfitting and enhance the reliability of model evaluation due to the limited sample size. This research used six unique machine learning algorithms: Logistic Regression, Random Forest, Extreme Gradient Boosting, Light Gradient Boosting Machine, Support Vector Machine, and Artificial Neural Network. Grid-search cross-validation was used to refine the hyperparameters. Models were eventually chosen based on cross-validation results and then validated using the independent testing dataset.

The main result was in-hospital mortality, defined as death from any reason during the first hospitalization. Individuals discharged alive were classified as survivors. The secondary outcomes were ICU admission, need for mechanical breathing, provision of inotropic support, occurrence of cardiogenic shock, and duration of hospitalization. The information was examined using conventional statistical methods. Continuous variables were reported as mean ± standard deviation or median (interquartile range) according on their distribution, whilst categorical variables were conveyed as counts and percentages. An independent-samples t-test or Mann–Whitney U test was used to analyze continuous data, whilst a chi-squared test or Fisher's exact test was applied to categorical variables comparing survivors and non-survivors.

The effectiveness of traditional risk assessments and machine-learning algorithms was assessed by calculating the area under the receiver operating characteristic curve, in addition to metrics such as accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1-score. Calibration plots and the Hosmer–Lemeshow test for goodness-of-fit were used to assess calibration where suitable. Statistically significant results were defined by a p-value below 0.05.

The research obtained authorization from the Institutional Review Board of Sheikh Zayed Medical College, Rahim Yar Khan, Pakistan (Reference No. SZMC/IRB/2024/127). All individuals involved or their legally designated representatives submitted signed informed consent. Patient confidentiality was preserved by

the anonymization of all gathered data, and the research was executed in accordance with the tenets of the Declaration of Helsinki.

RESULTS

There were 120 acute heart failure patients included in the current study. The mean age of the study group was 63.8 ± 11.4 years. There were 70 males (58.3%) and 50 females (41.7%). 18 patients (15.0%) died in the hospital, and 102 patients (85.0%) were discharged alive. Non-survivors were older and were more likely than survivors to have diabetes mellitus, chronic kidney disease, atrial fibrillation, hypotension, hypoxemia, renal dysfunction, hyponatremia, elevated lactate level, and lower left ventricular ejection fraction, as well as cardiogenic shock and inotropic support (Table 1). Moderate in-hospital mortality predictive capacity was obtained for each of the traditional clinical risk scores. Of these, the GWTG-HF score had the highest AUC value of 0.78, followed by the ADHERE risk tree (AUC 0.75) and the OPTIMIZE-HF score (AUC 0.73). Their sensitivity, on the other hand, was relatively low, which shows that conventional scoring systems were not able to correctly identify some high-risk patients (Table 2).

Machine-learning models demonstrated better predictive performance than traditional risk scores. The best performance was obtained by XGBoost (AUC: 0.91, accuracy: 88.3%, sensitivity: 83.3%, specificity: 89.2%, F1-score: 0.76). LightGBM had the second largest AUC of 0.89, followed by Random Forest (AUC 0.87) and artificial neural network (AUC 0.86). Logistic regression and support vector machine also showed better performance than traditional scores, but their accuracy is still lower than the models based on ensemble (Table 2). According to the XGBoost model, SBP was the best predictor of in-hospital mortality, followed by serum creatinine, lactate level, oxygen saturation, LVEF, serum sodium, inotropic support, age, chronic kidney disease, and heart rate. These predictors represented the hemodynamic, renal, tissue perfusion, hypoxemia, and poor cardiac function factors contributing to early death in acute heart failure (Table 3).

Comparative AUC performance of traditional risk scores and machine-learning models is shown in Figure 1. The AUC values of traditional risk scores were below 0.80, while most of the AUC values of machine-learning models were > 0.80. XGBoost showed the greatest AUC, followed by LightGBM and Random Forest models; thus, the discriminatory power of the machine-learning models was the best among the three models in predicting in-hospital mortality in acute heart failure. Machine-learning models, particularly XGBoost, outperformed traditional risk scores and more accurately identified patients at high risk of in-hospital mortality.

Table 1: Baseline clinical and laboratory characteristics of survivors and non-survivors

Variable	Survivors (n=102)	Non-survivors (n=18)	p-value
Age, years	62.4 ± 10.8	71.7 ± 9.6	0.001
Male	58 (56.9%)	12 (66.7%)	0.439
Female	44 (43.1%)	6 (33.3%)	0.439
Diabetes mellitus	48 (47.1%)	13 (72.2%)	0.049
Hypertension	66 (64.7%)	14 (77.8%)	0.279
Chronic kidney disease	25 (24.5%)	10 (55.6%)	0.008
Ischemic heart disease	54 (52.9%)	13 (72.2%)	0.132
Atrial fibrillation	21 (20.6%)	8 (44.4%)	0.029
Systolic blood pressure, mmHg	126.8 ± 22.4	96.6 ± 18.7	<0.001
Diastolic blood pressure, mmHg	78.4 ± 12.6	62.8 ± 10.5	<0.001
Heart rate, beats/min	104.2 ± 18.1	121.9 ± 20.4	0.001
Respiratory rate/min	24.6 ± 5.2	31.8 ± 6.1	<0.001
Oxygen saturation, %	91.9 ± 5.4	84.8 ± 6.6	<0.001
Serum sodium, mmol/L	136.3 ± 4.7	130.5 ± 5.2	<0.001
Serum creatinine, mg/dL	1.43 ± 0.61	2.32 ± 0.88	<0.001
Blood urea nitrogen, mg/dL	38.6 ± 14.8	67.4 ± 22.5	<0.001
Lactate, mmol/L	2.1 ± 0.8	4.4 ± 1.4	<0.001
Left ventricular ejection fraction, %	37.9 ± 9.5	29.3 ± 8.4	0.001
Inotropic support	17 (16.7%)	13 (72.2%)	<0.001
Cardiogenic shock	8 (7.8%)	11 (61.1%)	<0.001

Table 2: Predictive performance of traditional risk scores and machine-learning models

Model	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1-score
ADHERE risk tree	0.75	76.7%	61.1%	79.4%	34.4%	92.0%	0.52
GWTG-HF score	0.78	79.2%	66.7%	81.4%	38.7%	93.3%	0.57
OPTIMIZE-HF score	0.73	75.0%	55.6%	78.4%	31.3%	90.9%	0.48
Logistic regression	0.82	81.7%	72.2%	83.3%	43.3%	94.4%	0.62
Support vector machine	0.84	83.3%	72.2%	85.3%	46.4%	94.6%	0.65
Random Forest	0.87	85.8%	77.8%	87.3%	51.9%	95.7%	0.70
LightGBM	0.89	87.5%	83.3%	88.2%	55.6%	96.8%	0.74
XGBoost	0.91	88.3%	83.3%	89.2%	57.7%	96.8%	0.76
Artificial neural network	0.86	84.2%	77.8%	85.3%	48.3%	95.6%	0.68

Table 3: Top predictors of in-hospital mortality according to XGBoost feature importance

Rank	Predictor	Relative importance
1	Systolic blood pressure	0.162
2	Serum creatinine	0.141
3	Lactate level	0.128
4	Oxygen saturation	0.113
5	Left ventricular ejection fraction	0.096
6	Serum sodium	0.084
7	Inotropic support	0.079
8	Age	0.071
9	Chronic kidney disease	0.066
10	Heart rate	0.060

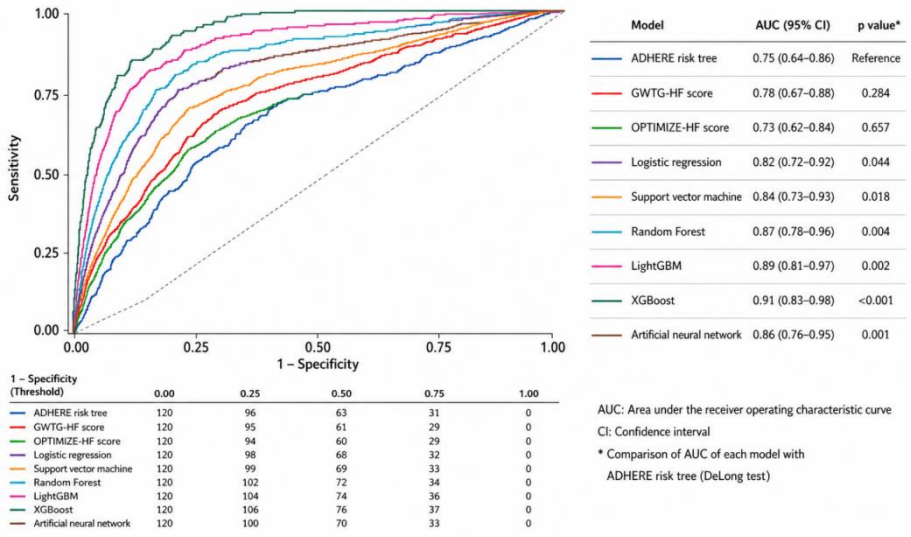

Figure 1: Comparative AUC performance of traditional risk scores and machine-learning models

DISCUSSION

The present study compared machine-learning models with traditional clinical risk scores for predicting in-hospital mortality among patients admitted with acute heart failure [1]. The results showed that machine-learning models outperformed traditional bedside scoring systems. XGBoost had the best discrimination, accuracy, sensitivity, specificity, and F1-score among all the models evaluated, followed by LightGBM and Random Forest. Moderate predictive power was noted with traditional risk scores such as ADHERE, GWTG-HF and OPTIMIZE-HF, but these were less accurate than machine-learning strategies [2,3].

This research reported an in-hospital death rate of 15.0%, indicating a high-risk patient group [4]. Non-survivors exhibited more severe clinical characteristics at presentation, including advanced age, reduced systolic and diastolic blood pressure, elevated heart rate, increased respiratory rate, diminished oxygen saturation, renal impairment, hyponatremia, elevated lactate levels, decreased left ventricular ejection fraction, cardiogenic shock, and a greater requirement for inotropic support [5,6].

The results of these studies suggest that hemodynamic instability, inadequate tissue perfusion, renal dysfunction, hypoxia, and severe myocardial dysfunction are important factors in early mortality in acute heart failure [7].

Traditional risk scores are clinically valuable because they are simple and quick and utilize parameters that are routinely available. GWTG-HF outperformed ADHERE and OPTIMIZE-HF in the current study, and each of the conventional scores was less sensitive than the machine-learning models [1-8]. This indicates that conventional approaches could underestimate a fraction of patients at high risk for mortality, particularly if they have more than one abnormality that interacts. Their low performance may be due to the use of a small number of variables and insufficiently considering the possible nonlinear interactions between clinical, laboratory, and echocardiographic predictors [9-11].

Machine-learning models outperformed because they are able to handle a wider variety of variables and to detect more intricate relationships between predictors and outcomes [12]. The model with the best performance was XGBoost with an AUC of 0.91, which has excellent discrimination. Its superior performance could be due to its ability to capture nonlinear associations and interactions among variables more effectively, as well as imbalanced patterns of outcomes, than conventional statistical models.

The performance of LightGBM and Random Forest was also good, demonstrating that ensemble-based algorithms are a valuable tool for the prediction of acute cardiovascular risk [13].

Systolic blood pressure, serum creatinine, lactate level, oxygen saturation, left ventricular ejection fraction, serum sodium, inotropic support, age, chronic kidney disease, and heart rate were the most important predictors identified by the XGBoost model. These variables are clinically meaningful and biologically plausible [14]. Low systolic blood pressure indicates low cardiac output and circulatory compromise. Cardiorenal dysfunction can be identified by elevated serum creatinine or chronic kidney disease, and is associated with poor prognosis in acute heart failure. High lactate levels indicate inadequate tissue perfusion, and low oxygen saturation indicates a problem with breathing and congestion in the lungs. Low ejection fraction and the need for inotropic support are signs of severe cardiac dysfunction and instability [15].

A major strength of the current study is that it was of a prospective nature. Data collection was done at the time of index admission to avoid incomplete or incorrect retrospective documentation. One of the assets is the direct comparison of several ML models to current traditional risk scores in the same patient population. This enabled more clinically relevant assessment of the predictive ability. The most frequently used clinical and laboratory variables are also included, increasing the relevance of the results in everyday hospital practice [16].

The results are of clinical relevance. Identification of high-risk acute heart failure patients early in the process may aid in prioritizing intensive monitoring, early cardiology review, early ICU referral, optimization of diuretics and vasodilators, timely use of inotropes as indicated, and closer follow-up after stabilization [2,17]. The machine-learning model can be integrated into electronic medical record systems as an automated decision support system to provide an estimate of death risk at time of admission. They may be helpful in addition to, but not in place of, clinical judgment [5-8,18].

There are some downsides to it, however. First, the patients were few in number (120) and the mortality rate in the hospital was low (18), which makes it difficult to obtain a stable result in the model and is easy to overfit the model. A cross validation was carried out, and other multicenter data sets are required for additional validation. Second, the study was carried out for a short period of time, and external validation was not done. Third, some advanced biomarkers and serial measurements can be performed dynamically, which can further improve the model's performance. Fourth, machine-learning models require technical infrastructure and familiarity with the models by the clinician before they can be implemented routinely [19,20].

However, the findings of this study reveal that machine learning models, especially the XGBoost model, might be superior to traditional risk scores for predicting in-hospital mortality in acute heart failure. The results suggest that machine-learning clinical decision-support systems could be further developed and validated in larger and more diverse patient cohorts [17-20].

CONCLUSION

Machine-learning models performed better than traditional risk scores to predict in-hospital mortality in patients hospitalized with acute heart failure. Among the three models, XGBoost performed best followed by LightGBM and Random Forest. The traditional risk scores also proved helpful, although with less discrimination and sensitivity. Adverse factors that were the best predictors of death included hypotension, renal dysfunction, elevated lactate, hypoxemia, reduced left ventricular ejection fraction, hyponatremia, inotropic requirement, advanced age, chronic kidney disease, and tachycardia. These results indicate that machine learning models could offer more accurate early risk stratification and could aid in identifying patients who are high risk for whom monitoring should be more frequent, and care should be escalated sooner.

Conflict of Interest: The authors declare that they have no conflict of interest.

Funding: This research received no external funding.

Authors' Contributions: A.H. and S.Y. conceived and designed the study. H.A. performed data analysis, machine-learning model development, and manuscript drafting. M.K. contributed to data interpretation, critical revision of the manuscript, and overall supervision. All authors reviewed and approved the final version of the manuscript.

Acknowledgments: The authors sincerely acknowledge the support of Sheikh Zayed Medical College, Rahim Yar Khan, Pakistan, and the University of Hertfordshire, Hatfield, United Kingdom, for facilitating this research.

Data Availability: The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

REFERENCES:

1. Sun M, Lu P, Cheng J, Zhao L, Ma J. Early prediction of in-hospital mortality in ICU patients with heart failure using an interpretable machine learning model: a retrospective study based on MIMIC-IV. *BMC Cardiovasc Disord.* 2026;26:552. doi:10.1186/s12872-026-05943-1.
2. Su D, Zheng J, Shao YK, Liu JY, Liu XX, Yu K, et al. Developing and validating a machine learning-based model for predicting in-hospital mortality among ICU-admitted heart failure patients: a study utilizing the MIMIC-III database. *Digit Health.* 2025;11:20552076251335705. doi:10.1177/20552076251335705.
3. Li J, Sun Y, Ren J, Wu Y, He Z. Machine learning for in-hospital mortality prediction in critically ill patients with acute heart failure: a retrospective analysis based on the MIMIC-IV database. *J Cardiothorac Vasc Anesth.* 2025;39(3):666-674. doi:10.1053/j.jvca.2024.12.016.
4. Yan L, Zhang J, Chen L, Zhu Z, Sheng X, Zheng G, et al. Predictive value of machine learning for the risk of in-hospital

death in patients with heart failure: a systematic review and meta-analysis. *Clin Cardiol.* 2025;48(1):e70071. doi:10.1002/clc.70071.
5. Scrutinio D, Amitrano F, Guida P, Coccia A, Pagano G, D'Addio G, et al. Prediction of mortality in heart failure by machine learning: comparison with statistical modeling. *Eur J Intern Med.* 2025;133:106-112. doi:10.1016/j.ejim.2025.01.020.
6. Liu X, Xie Z, Zhang Y, Huang J, Kuang L, Li X, et al. Machine learning for predicting in-hospital mortality in elderly patients with heart failure combined with hypertension: a multicenter retrospective study. *Cardiovasc Diabetol.* 2024;23(1):407. doi:10.1186/s12933-024-02503-9.
7. Jawadi Z, He R, Srivastava PK, Fonarow GC, Khalil SO, Krishnan S, et al. Predicting in-hospital mortality among patients admitted with a diagnosis of heart failure: a machine learning approach. *ESC Heart Fail.* 2024;11(5):2490-2498. doi:10.1002/ehf2.14796.
8. Tanaka M, Kohjitani H, Yamamoto E, Morimoto T, Kato T, Yaku H, et al. Development of interpretable machine learning models to predict in-hospital prognosis of acute heart failure patients. *ESC Heart Fail.* 2024;11(5):2798-2812. doi:10.1002/ehf2.14834.
9. Huang J, Cai Y, Wu X, Huang X, Liu J, Hu D. Prediction of mortality events of patients with acute heart failure in intensive care unit based on deep neural network. *Comput Methods Programs Biomed.* 2024;256:108403. doi:10.1016/j.cmpb.2024.108403.
10. Li X, Wang Z, Zhao W, Shi R, Zhu Y, Pan H, et al. Machine learning algorithm for predicting in-hospital mortality in critically ill patients with congestive heart failure combined with chronic kidney disease. *Ren Fail.* 2024;46(1):2315298. doi:10.1080/0886022X.2024.2315298.
11. König S, Pellissier V, Hohenstein S, Leiner J, Meier-Hellmann A, Kuhlen R, et al. From population- to patient-based prediction of in-hospital mortality in heart failure using machine learning. *Eur Heart J Digit Health.* 2022;3(2):307-310. doi:10.1093/ehjdh/ztac012.
12. Li J, Liu S, Hu Y, Zhu L, Mao Y, Liu J. Predicting mortality in intensive care unit patients with heart failure using an interpretable machine learning model: retrospective cohort study. *J Med Internet Res.* 2022;24(8):e38082. doi:10.2196/38082.
13. Peng S, Huang J, Liu X, Deng J, Sun C, Tang J, et al. Interpretable machine learning for 28-day all-cause in-hospital mortality prediction in critically ill patients with heart failure combined with hypertension: a retrospective cohort study based on MIMIC-IV and eICU databases. *Front Cardiovasc Med.* 2022;9:994359. doi:10.3389/fcvm.2022.994359.
14. Li L, Zhang Z, Xiong Y, Hu Z, Liu S, Tu B, et al. Prediction of hospital mortality in mechanically ventilated patients with congestive heart failure using machine learning approaches. *Int J Cardiol.* 2022;358:59-64. doi:10.1016/j.ijcard.2022.04.063.
15. Park J, Hwang IC, Yoon YE, Park JB, Park JH, Cho GY. Predicting long-term mortality in patients with acute heart failure by using machine learning. *J Card Fail.* 2022;28(7):1078-1087. doi:10.1016/j.cardfail.2022.02.012.
16. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation.* 2022;145(18):e895-e1032. doi:10.1161/CIR.0000000000001063.
17. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J.* 2021;42(36):3599-3726. doi:10.1093/eurheartj/ehab368.
18. Peterson PN, Rumsfeld JS, Liang L, Albert NM, Hernandez AF, Peterson ED, 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.
19. Abraham WT, Fonarow GC, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, 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.
20. Fonarow GC, Adams KF Jr, Abraham WT, Yancy CW, Boscardin WJ; ADHERE Scientific Advisory Committee, Study Group, and Investigators. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *JAMA.* 2005;293(5):572-580. doi:10.1001/jama.293.5.572.

This Article May be cited as: Rahim AH, Yousaf S, Akram H, Kashif M. Machine learning versus traditional risk scores for predicting in-hospital mortality in acute heart failure: machine learning for acute heart failure mortality. *Dev Med Life Sci.* 2026;3(6):32-38. doi:10.69750/dmls.03.06.0217.

Publisher's Note:
Developmental Medico-Life-Sciences remains neutral with regard to jurisdictional claims in published maps. and institutional affiliations.

DMLS Publisher
Developmental Medico-Life-Sciences Research and Publications Pvt Ltd.