Research Article ISSN 3066-3830

International Journal of Forensic Science & Research

A Retrospective Cohort Study: Predictors of Mortality in Heart Failure

Edinen Asuka*, Islam Salem and Mathivanan Kamaraj

Yale school of Public Health, Connecticut, USA.

*Correspondence:
Edinen Asuka, Yale school of Public Health, Connecticut, USA, Phone: +13477200862.

Received: 15 Oct 2025; **Accepted:** 18 Oct 2025; **Published:** 29 Nov 2025

Citation: Edinen Asuka, Islam Salem, Mathivanan Kamaraj. A Retrospective Cohort Study: Predictors of Mortality in Heart Failure. Int J Forens Sci Res. 2025; 2(1): 1-4.

ABSTRACT

Aim: *Our objective in this study was to investigate the association between the variables: age, diabetes mellitus, ejection fraction, Hypertension and the outcome (mortality) in patients with heart failure and identify which of these are significant predictors of mortality.*

Methodology: *This study investigated clinical predictors of mortality in patients with heart failure (reduced or preserved ejection fraction) using a retrospective dataset containing records for 299 patients from Pakistan. The mean age was 60.8 years, and 32% of patients experienced a death event. A multivariable logistic regression model was fitted to identify significant predictors of mortality, holding other independent variables included in the model fixed.*

Results: *A multivariable logistic regression analysis revealed that age and ejection fraction were significant predictors of mortality. Increasing age was associated with higher odds of a death event (OR = 1.061, 95% CI [1.04–1.09], $p < 0.001$), holding other independent variables in the model fixed. Conversely, log-transformed ejection fraction was strongly protective, with each unit increase significantly reducing the odds of death (OR = 0.07, 95% CI [0.03–0.17], $p < 0.001$), holding other independent variables in the model fixed. Hypertension with $p = 0.266$ and diabetes mellitus ($p = 0.462$) were not statistically significant predictors of heart failure in this regression model.*

Conclusion: *The outcome of this study shows that age and ejection fraction are significant predictors of mortality in patients with heart failure (reduced or preserved ejection fraction). These findings underscore the importance of considering these clinical parameters for risk stratification in patient management.*

Keywords
Heart Failure, Mortality Predictors, Ejection Fraction, Age, Logistic Regression.

Introduction
Heart failure (HF) is a condition that affects the heart's inability to pump sufficient blood to meet the body's metabolic needs or it may also be due to diastolic dysfunction. It is one of the leading causes of hospitalization and mortality worldwide. This condition has mainly classified into heart failure with reduced ejection fraction (HFrEF) and heart failure with preserved ejection fraction. In heart failure with reduced ejection fraction, the heart muscle is weakened and cannot contract effectively while in heart failure with preserved ejection fraction (HFpEF), the heart muscle is stiff and cannot relax to fill with blood properly [1-3]. Among the numerous risk factors associated with heart failure, age and diabetes mellitus are among the major contributors to increased morbidity and mortality in both HFrEF and HFpEF [3-5]. There are an estimated 56 million people with heart failure (HF) worldwide. The prevalence of HF differs geographically [4,5]. According to Centers for Disease Control and Prevention (CDC), in the United States, nearly 6.7 million adults 20 years old or older have heart

failure. In 2023, heart failure was mentioned on 452,573 death certificates (approximately 14.6% of all-cause mortality) [4,5]. The American Heart Association (AHA) estimated that there were 6.7 million people (2.3 %) living with heart failure in the United States between 2017 and 2020 [4,5]. HFpEF appears to be associated with a better prognosis than HFrEF in some reports, although some investigators have reported similar mortality rates for HFrEF and HFpEF [4,5]. A meta-analysis by Global Group in Chronic Heart Failure (MAGGIC) which included about 41,972 patients with HF (31 studies) found that patients with HFpEF had a substantially lower risk of death compared with those with HFrEF. Risk factors or conditions associated with heart failure include diabetes mellitus, obesity, coronary artery disease, hypertension, age, alcohol, illicit substance use, certain medications, and smoking [4,5]. Advancing age is one of the risk factors for heart failure, leading to physiological changes in the cardiovascular system [5,6]. Chronic hyperglycemia due to diabetes mellitus has shown to contribute to microvascular and macrovascular damage. Diabetes mellitus can worsen existing coronary artery disease and lead to further declines in function. These factors create detrimental physiological changes that can adversely affect patients. Here, we will further explore the association of these risk factors and the outcome death in these patients. We will assess their role as independent predictors of death in patients with heart failure.

Methodology

This was a retrospective cohort study and the dataset [7] was obtained from Health Data Hub. Dataset used in this research included medical records of patients with heart failure (HFpEF and HFrEF), admitted at the Institute of Cardiology and Allied hospital Faisalabad-Pakistan, between April–December 2015. It consisted of 105 women and 194 men, with age range between 40 and 95 years old. All 299 patients had heart failure, with the minimum and maximum ejection fraction reported respectively, as 14% and 80%. The dataset consisted of 13 variables and 299 observations. The key variables used in this research included age (in years), diabetes mellitus (Yes/No), and death (Yes/No) as outcome. Other variables included in the multivariable regression analysis included Hypertension as Yes/No and ejection fraction (%). We used odds ratio as a measure of association. We also did a multivariable logistic regression to explore the association between the independent or predictor variables and the outcome (death). From the dataset and as outlined in Table 1, we were able to ascertain that there were 194 males (64.88%) and 105 females (35.12%). Overall, the mean age of the patients was 60.8 years. Males had a slightly higher mean age of 61.4 years with 95% confidence interval [59.7, 63.1] compared to females at 59.78 years with 95% confidence interval [57.6, 62.0] with a p-value of 0.2478, which indicates no statistically significant difference in age between both groups. A total of 96 patients (32%) experienced a death event. The majority of patients were over 50 years of age. More details about characteristics of study participants outlined in Table 2.

Table 1: Summary of demographics.

Characteristic	Statistic	Value
Total Patients	N	**299**
Sex Distribution		
Male	N (%)	194 (64.88%)
Female	N (%)	105 (35.12%)
Age Statistics		
Overall Mean Age (years)	Mean	60.8
Male Mean Age (95% CI)	Mean (95% CI)	61.4 [59.7, 63.1]
Female Mean Age (95% CI)	Mean (95% CI)	59.78 [57.6, 62.0]
Age Difference (Male vs Female)	P-value	0.2478
Primary Outcome		
Death Event	N (%)	96 (32.11%)

Note: CI = Confidence interval.

Table 2: Characteristics of Study Participants.

Characteristic	No Death (N = 203)	Death (N = 96)	P-value
Total Sample Size, N	203	96	---
Age (years)*	58.8 ± 10.6	65.2 ± 13.2	< 0.001
Creatinine phosphokinase (units/L)*	540 ± 754	670 ± 1317	0.369
Ejection fraction (%)*	40 ± 11	33 ± 13	< 0.001
Platelets (kiloplatelets/mL)*	266, 657 ± 97, 531	256, 381 ± 98, 526	0.399
Serum creatinine (mg/dL)*	1.18 ± 0.65	1.84 ± 1.47	< 0.001
Serum sodium (mEq/L)*	137.2 ± 4.0	135.4 ± 5.0	0.002
Follow-up time (days)*	158 ± 68	71 ± 62	< 0.001
Anemia (Yes)	83 (40.9%)	46 (47.9%)	0.307
Diabetes mellitus (Yes)	85 (41.9%)	40 (41.7%)	1.000
Hypertension (Yes)	66 (32.5%)	39 (40.6%)	0.214
Male sex	132 (65.0%)	62 (64.6%)	1.000
Female sex	71 (35.0%)	34 (35.4%)	1.000
Smoking history (Yes)	66 (32.5%)	30 (31.2%)	0.932

*(Mean ± Standard deviation)

Statistical Methods

Descriptive statistics was utilized to explore the characteristics of study participants. A multivariable logistic regression model was fitted to identify significant predictors of mortality in these patients. The model included age, ejection fraction, diabetes mellitus, and hypertension as predictor or independent variables. Death was the dependent variable, binary outcome (Yes/No). The model's coefficients were converted to provide odds ratios with 95% confidence intervals. The overall significance of the model (global hypothesis test) was checked using the Likelihood Ratio Test. We carried out the variance inflation test to check for the presence of multicollinearity. We used Box-Tidwell transformation terms with the likelihood ratio test to assess the presence of the linearity of the logit assumption. R version 4.5.1 was used for statistical analysis and data visualization.

Logistic regression: Assumptions and implications

Binary or categorical outcome: The dependent or outcome variable must be a categorical or binary outcome. If the outcome has more than two categories, you cannot use standard binary logistic regression; a multinomial logistic regression (for nominal outcomes with more than two categories) or ordinal logistic regression (for ordinal outcomes with more than two categories) can be used. Here, our outcome, death, had two categories (death or no death), so it was ideal to use a logistic regression model here.

Independence of Observations: The observations must be independent of each other. This means that there is no relationship between the outcomes of different individuals or cases. A lack of independence can lead to incorrect standard errors, making the p-values unreliable. This can increase the likelihood of a Type I error. We felt this assumption was not violated based on the data used in this study.

Linearity of Log-Odds: There must be a linear relationship between the continuous predictor variables and the logit (the natural log of the odds). If this assumption is violated, the model's coefficients may be biased, and the model may not accurately predict the outcome. This can lead to a poor fit and a misleading interpretation of the odds ratios or less precise model. Data on diabetes mellitus and hypertension in this study were in two categories (not as continuous numeric forms), so we were unable to assess linearity of logit for these variables. We were able to establish that for the variable age, the linearity of logit was met and for the variable ejection fraction, the linearity of logit was not met. On that note, the variable ejection fraction was log-transformed.

No Multicollinearity: Predictor variables should not be highly correlated with each other. This is known as multicollinearity. High multicollinearity makes it difficult for the model to estimate the unique or independent effect of each predictor, which can lead to large standard errors and unstable, unreliable coefficients. We assessed the presence of multicollinearity by using Variance Inflation Factor (VIF) and this assumption was met. However, if there was any significant violation, we would aim to address it by removing one of the highly correlated variables or combining them.

Large Sample Size: Logistic regression analysis as in this case requires sufficient sample size to provide stable and reliable results. A common requirement is to have at least 10 events per predictor variable. A small sample size can lead to unstable and biased coefficient estimates, and the model's predictions may not generalize well to the larger population. It can also result in overtly wide confidence intervals, making the results less precise.

Multivariable logistic regression results

Ejection fraction (log-transformed): Coefficient (β): −2.73 (Negative); Odds Ratio (OR): 0.07, p value <0.001. The odds ratio of 0.07 corresponds to a 1-unit increase in natural log(EF). And because log(EF) is the natural log, a 1-unit increase in log(EF) means multiplying EF itself by e (~2.72). If a patient's ejection fraction (EF) goes up by about 2.7 times, their odds of dying drop by about 93%, holding other independent variables included in the model fixed.

Age: Coefficient (β): 0.06 (Positive); Odds Ratio (OR): 1.06, P-value <0.001. The odds ratio is greater than 1. For every one-year increase in age, the odds of a death event increase by approximately 6.0%, holding other independent variables included in the model fixed. The 95% CI (1.04 to 1.09).

Hypertension: Coefficient (β): 0.32 (Positive), Odds Ratio (OR): 1.37, 95% CI (0.79 to 2.40). P-value: 0.266. The odds of death for a patient with Hypertension are 37% higher compared to one without, this effect is not statistically significant holding other independent variables included in the model fixed.

Diabetes mellitus: Coefficient (β): 0.21 (Positive), Odds Ratio (OR): 1.23, 95% CI (0.71 to 2.14), P-value: 0.462. The odds of death for a patient with diabetes are 23% higher compared to one without, but this effect is also not statistically significant holding other independent variables included in the model fixed.

Therefore, based on the result of this multivariable logistic regression model as outlined in Table 3, the variables age and ejection fraction are statistically significant predictors of the death in patients with heart failure. While diabetes mellitus and Hypertension though appear to be associated with an increase in the odds of the outcome, are not statistically significant.

Table 3: Summary of logistic regression model.

Term	estimate	std. error	statistic	p.value	conf. low	conf. high
(Intercept)	157.00	1.68	3.02	<0.05	6.22	4581.00
log(ejection_ fraction)	0.07	0.48	-5.67	<0.001	0.02	0.16
Age	1.06	0.01	4.79	<0.001	1.04	1.09
Hypertension	1.37	0.28	1.11	0.27	0.78	2.40
Diabetes mellitus	1.23	0.28	0.74	0.46	0.71	2.14

Discussion

The identification of age and ejection fraction as statistically significant predictors of mortality underscores their crucial roles in the risk assessment of these patients. An increasing age is a well-known risk factor for numerous health complications, and in heart failure, it often correlates with a more advanced disease state and reduced physiological reserves [8,9]. Likewise, ejection fraction, a key measure of the heart's pumping efficiency, is a primary indicator of cardiac function [10,11].

A lower ejection fraction suggests a greater impairment in the heart's ability to circulate blood, directly increasing the risk of adverse events and mortality. Although, while diabetes mellitus and hypertension are known comorbidities that can worsen cardiovascular outcomes, they did not emerge as statistically significant predictors of mortality in this specific multivariable model. This result does not negate their clinical importance but may be attributed to the relatively small sample size or the specific characteristics of the patient population in this study. It is also possible that the effects of these independent variables were confounded by other possible variables not included in this model.

With regards to the limitations of this study, the study design used here prevents the definitive establishment of causality. The findings in this study are applicable to patients with heart failure, reduced or preserved ejection fraction. As earlier mentioned, we were able to establish the assumption of linearity of logit for the independent or predictor variables, age and ejection fraction using Box-Tidwell transformation terms with likelihood ratio test. Based on the outcome of the above test, we decided to carry out a logarithmic transformation of the variable, ejection fraction since it did not meet this assumption.

Future Research

We encourage use of a larger sample size in subsequent research to increase the power of the study. Likewise, we recommend use of time-event analysis models such as the Kaplan Meier model or Cox proportional hazards model in future research on this subject, as this would add more information to the subject matter. Also, to gain more information, we recommend exploration of other independent variables like smoking, and presence or history of ischemic heart disease. Variables should be clearly defined as well to establish more clarity. Future research could benefit from larger, multi-center studies to confirm these findings and explore the long-term impact of these variables on patient outcomes.

Conclusion

Based on the findings in this study, age and ejection fraction are significant predictors of mortality in patients with heart failure (reduced or preserved ejection fraction). Therefore, it is imperative that clinicians consider these factors when risk stratifying patients with heart failure.

References

1. Lloyd-Jones DM, Larson MG, Leip EP, et al. Framingham Heart Study. Lifetime risk for developing congestive heart failure: the Framingham Heart Study. Circulation. 2002; 106: 3068-3072.
2. https://www.acc.org/Latest-in-Cardiology/Articles/2021/07/12/12/31/Universal-Definition-and-Classification-of-Heart-Failure
3. https://www.ncbi.nlm.nih.gov/books/NBK430873/
4. https://www.uptodate.com/contents/epidemiology-of-heart-failure/print
5. Berry C, Doughty RN, Granger C, et al. Meta-analysis Global Group in Chronic Heart Failure (MAGGIC). The survival of patients with heart failure with preserved or reduced left ventricular ejection fraction: an individual patient data meta-analysis. Eur Heart J. 2012; 33: 1750-1757
6. https://www.cdc.gov/heart-disease/about/heart-failure.html
7. https://data.poltekkes-smg.ac.id/dataset/heart-failure-prediction
8. https://www.heart.org/en/health-topics/heart-failure/causes-and-risks-for-heart-failure
9. Rodgers JL, Jones J, Bolleddu SI, et al. Cardiovascular Risks Associated with Gender and Aging. J Cardiovasc Dev Dis. 2019; 6: 19.
10. Solomon SD, Anavekar N, Skali H, et al. Candesartan in Heart Failure Reduction in Mortality (CHARM) Investigators. Influence of ejection fraction on cardiovascular outcomes in a broad spectrum of heart failure patients. Circulation. 2005; 112: 3738-3744.
11. Chavey WE, Hogikyan RV, Van Harrison R, et al. Heart Failure Due to Reduced Ejection Fraction: Medical Management. Am Fam Physician. 2017; 95: 13-20.