International Journal of General Medicine

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Impact of Diabetes Mellitus on Heart Failure Patients: Insights from a Comprehensive Analysis and Machine Learning Model Using the Jordanian Heart Failure Registry

Mahmoud Izraiq[1], Eyas Almousa[2], Suhail Hammoudeh[1], Mazen Sudqi[1], Yaman B Ahmed[3], Omran A Abu-Dhaim[1], Abdel-Latif Mughrabi Sabbagh[1], Karam I Khraim[1], Ahmad A Toubasi[4], Abdullah Al-Kasasbeh[3], Sukaina Rawashdeh[3], Hadi Abu-Hantash[5]

[1]Cardiology Section, Internal Medicine Department, Specialty Hospital, Amman, Jordan; [2]Department of Cardiology, Istishari Hospital, Amman, Jordan; [3]Cardiology Section, Internal Medicine Department, King Abdullah University Hospital, Irbid, Jordan; [4]Cardiology Section, Internal Medicine Department, Jordan University Hospital, Amman, Jordan; [5]Department of Cardiology, Amman Surgical Hospital, Amman, Jordan

Correspondence: Mahmoud Izraiq, Cardiology Section, Internal Medicine Department, Specialty Hospital, Amman, Jordan, Tel +962795652260, Email izraiq@yahoo.com

Background: Heart failure (HF) is a common final pathway of various insults to the heart, primarily from risk factors including diabetes mellitus (DM) type 2. This study analyzed the clinical characteristics of HF in a Jordanian population with a particular emphasis on the relationship between DM and HF.

Methods: This prospective study used the Jordanian Heart Failure Registry (JoHFR) data. Patients with HF were characterized by DM status and HF type: HF with preserved ejection fraction (HFpEF) or HF with reduced ejection fraction (HFrEF). Demographics, clinical presentations, and treatment outcomes were collected. Statistical analyses and machine learning techniques were carried out for the prediction of mortality among HF patients: Recursive Feature Elimination with Cross-Validation (RFECV) and Synthetic Minority Over-sampling Technique with Edited Nearest Neighbors (SMOTEENN) were employed.

Results: A total of 2007 patients with HF were included. Notable differences between diabetic and non-diabetic patients are apparent. Diabetic patients were predominantly male, older, and obese ($p < 0.001$ for all). A higher incidence of HFpEF was observed in the diabetes cohort ($p = 0.006$). Also, diabetic patients had significantly higher levels of cholesterol ($p = 0.008$) and LDL ($p = 0.003$), reduced hemoglobin levels ($p < 0.001$), and more severe renal impairment (eGFR; $p = 0.006$). Machine learning models, particularly the Random Forest Classifier, highlighted its superiority in mortality prediction, with an accuracy of 90.02% and AUC of 80.51%. Predictors of mortality included creatinine levels >115 µmol/L, length of hospital stay, and need for mechanical ventilation.

Conclusion: This study underscores notable differences in clinical characteristics and outcomes between diabetic and non-diabetic heart failure patients in Jordan. Diabetic patients had higher prevalence of HFpEF and poorer health indicators such as elevated cholesterol, LDL, and impaired kidney function. High creatinine levels, longer hospital stays, and the need for mechanical ventilation were key predictors of mortality.

Keywords: heart failure, diabetes mellitus, Jordan, clinical characteristics, machine learning, mortality prediction, predictive analytics

Introduction

Heart failure (HF) is a complex clinical syndrome arising from any structural or functional cardiac disorder, and the structural base for heart failure development is systolic or diastolic myocardial dysfunction. It represents the end stage of various diseases affecting the heart's components and metabolic cardiovascular activities, mainly targeting the left ventricle.[1] Typical clinical picture of heart failure includes breathlessness, ankle swelling, fatigue, and signs such as elevated jugular venous pressure, pulmonary crackles, and peripheral edema. Heart failure can be distinguished because

Received: 29 March 2024
Accepted: 15 May 2024
Published: 18 May 2024

of ischemic or non-ischemic damage to the heart muscle. The risk factors for ischemic cardiomyopathy include typical risk factors for the development of atherosclerosis, such as diabetes, hypertension, dyslipidemia, nicotine addiction, obesity, and low physical activity. The progression to HF involves ventricular dilatation and remodeling, leading to decreased cardiac output and increased intracardiac pressures, affecting more than 26 million people globally. The risk factors for HF, such as hypertension, coronary artery disease, obesity, and type 2 diabetes mellitus (DM), either alone or in combination with dyslipidemia, hypertension, and obesity, are particularly pertinent.[2–4] DM precipitate and worsens the course of HF because of the buildup of advanced glycation end products, increased oxidative stress, compromised inflammatory responses, deterioration of intracellular calcium levels, alterations in microRNA expression, as well as the advancement of atherosclerosis and coronary artery disease.[5] Clinically, studies have demonstrated that patients with concomitant DM and HF have higher mortality.[6,7] Given the alarming rise in global and regional diabetes projections, understanding this relationship is crucial. Notably, prior research has predominantly focused on Western populations, with limited data available on the intersection of HF and DM in Middle Eastern settings, particularly in Jordan. Due to that, this study primarily aimed at comparing the clinical characteristics and outcomes of HF patients with and without diabetes mellitus (DM) in the Jordanian context. This investigation will provide vital insights into how DM influences the presentation and prognosis of HF in this specific population. Additionally, we aim to employ machine learning techniques to predict mortality among these patients, offering a novel approach to evaluating their clinical trajectories. Secondary objectives include analyzing the prevalence and impact of DM on hospitalization rates and lengths of stay among HF patients, evaluating the effect of DM on mortality rates and clinical complications, and assessing the healthcare utilization and economic burdens associated with HF in the context of DM. These objectives are crucial for developing targeted interventions that enhance patient care and manage the public health impact of heart failure more effectively in Jordan.

Methods

Study Design and Setting

Data for this study were obtained from the Jordanian Heart Failure Registry (JoHFR), which includes records of patients with acute and chronic heart failure seen in cardiology clinics and hospitals across Jordan from July 1st, 2021, to February 28th, 2023. This comprehensive registry facilitated longitudinal follow-up at 3, 6, 9, and 12 months to document changes in laboratory results, emergence of complications, and modifications to treatment plans.

Ethical Consideration

The study received ethical approval from each participating center's Institutional Review Board (IRB) and was registered at clinicaltrials.gov (NCT04829591). Following the ethical guidelines and standards outlined in the Declaration of Helsinki, we hereby confirm that our study fully complies with these principles. The Research Committee of the Faculty of Medicine and the Institutional Review Board at the Specialty Hospital approved the study, and the Institutional Review Board provided the ethical approval. The ethics committee approved a waiver of consent from the patients because the study did not include any therapeutic intervention and the outcomes planned are routinely registered in patients with heart failure.

Data Collection and Variables Measurement

An online form designed to collect patient data was designed for completion by healthcare professionals. The form was structured into 10 sections to capture a range of information, including medical history, heart failure status, laboratory tests and procedures performed by patients, treatment outcomes, mechanical ventilation requirement, hospital length of stay, morbidity, or mortality occurrence. The primary aim of data collection was to explore the outcomes in patients with heart failure overall and then in those with or without DM. Patients are further categorized, compared with and without DM, and further stratified as having heart failure with preserved ejection fraction (HFpEF) or heart failure with reduced ejection fraction (HFrEF). This allowed for a more granular analysis of the effect of DM on patient outcomes across the spectrum of HFpEF and HFrEF. Key variables collected included demographics (sex, age), clinical presentation, and HF

characteristics (BMI, current smoking status, alcohol use, symptoms [fatigue and dyspnea], type of HF [HFpEF vs HFrEF]), comorbidities (HTN, dyslipidemia, CKD, atrial fibrillation, ASCVD), laboratory findings, outcomes from treatments (need for mechanical ventilation, number of admissions, length of stay, 30-day mortality), and echo results (left ventricle ejection fraction (LV EF), pulmonary artery systolic pressure [PASP], left ventricular hypertrophy [LVH], and left atrial enlargement [LAE]).

Data Analysis

Initial data entry was conducted using Microsoft Office Excel 2019. The analysis of the data was performed using IBM SPSS version 25 (IBM Corp., Armonk, N.Y., USA). For continuous variables, central tendency and dispersion were assessed using the mean and standard deviation when data followed a normal distribution, as verified by Shapiro–Wilk tests. For non-normally distributed data, the median and interquartile ranges were used. Continuous variables were compared using *t*-tests, and categorical variables were analyzed using chi-square tests. A p-value of less than 0.05 was considered statistically significant. Missing data were handled by multiple imputations using the "mice" package in R. To ensure a robust analysis in the presence of potential data gaps, we created five imputed datasets using Predictive Mean Matching (PMM).

Machine Learning Analysis

Machine learning was employed to predict mortality in this study, utilizing Python 3.8 for all computational tasks on a MacBook equipped with an M1 processor and 16 GB RAM. For handling missing values, median imputation was used for numerical columns and the most frequent value for categorical columns, with data manipulation facilitated by the Pandas library, version 1.2.4. To ensure comparability of predictive features, all numerical data were standardized using the StandardScaler from the Scikit-learn library, version 0.24.2. Feature selection was conducted using Recursive Feature Elimination with Cross-Validation (RFECV) with a RandomForestClassifier, focusing on isolating the most relevant features for mortality prediction. The dataset was divided into training, validation, and test sets, with 40% reserved for validation and testing. For handling imbalanced datasets, the Synthetic Minority Over-sampling Technique and Edited Nearest Neighbors (SMOTEENN) were applied using the Imbalanced-learn library, version 0.8.0.[8] A grid search with five-fold cross-validation on the resampled training data was conducted to identify the optimal set of hyperparameters, after which the final model was trained. We compared the predictive capabilities of four different machine learning models: Random Forest Classifier (RFC),[9] Logistic Regression (LR),[10] Support Vector Machine (SVM),[11] and eXtreme Gradient Boosting (XGBoost).[12] These models were evaluated on their performance metrics such as accuracy, specificity, sensitivity, and Area Under the ROC Curve (AUC), employing numerical operations facilitated by the Numpy library, version 1.20.3. Permutation Feature Importance was utilized to assess the impact of each feature on our model's predictions by measuring changes in accuracy when feature values were randomly shuffled. This technique, integrated through Scikit-learn's feature importance functionality, provides an intuitive means to understand the relevance of each feature within our model's predictive framework, ensuring a comprehensive evaluation of the data's influence on patient outcomes.

Results

Clinical Characteristics and Laboratory Variables

This study analyzed the clinical characteristics and laboratory variables of heart failure patients, where the registry included 2151 patients, of whom 2007 met the inclusion criteria, including 1388 patients with diabetes mellitus (DM) and 619 patients without DM (No DM). This comparison revealed statistically significant differences in several clinical characteristics, Table 1. Notably, the prevalence of DM was associated with an increased incidence of males ($p < 0.001$), patients in the age group ≥70 years ($p < 0.001$), and those with a BMI categorization as obese ($p < 0.001$). The heart failure type was also significantly associated with DM, with a higher proportion of HFpEF observed in the DM group ($p = 0.006$). Mechanical ventilation and the number of hospital admissions were not significantly different between the groups, indicating similar acute care needs, irrespective of DM status. Laboratory findings highlighted the impact of DM

Table 1 Clinical Characteristics of Diabetic and Non-Diabetic Patients with Heart Failure

Variable, n (%)		DM (n=1388)	No DM (n=619)	p-value
Sex	Male	753 (39.8)	405 (43.6)	<0.001*
	Female	618 (60.2)	210 (56.4)	
Age	<40	21 (6.3)	41 (3.0)	<0.001*
	40–49	75 (10.5)	56 (5.8)	
	50–59	187 (19.3)	114 (14.3)	
	60–69	379 (27.0)	114 (26.1)	
	≥70	620 (36.8)	245 (50.8)	
Clinical Presentation	Fatigue	368	137	0.105
	Dyspnea	1067	445	0.267
	Orthopnea	520	179	0.001*
	Chest pain	425	167	0.430
	Palpitations	127	46	0.314
	PND	403	1232	0.006*
Hypertension	No	189 (20.6)	197 (19.0)	<0.001*
	Yes	1193 (79.4)	422 (81.0)	
BMI	Normal	182	119	<0.001*
	Overweight	271	155	
	Obese	350	89	
Smoking	No	1025 (32.4)	349 (29.4)	<0.001*
	Yes	357 (67.6)	270 (70.6)	
Alcohol	No	610 (59.5)	610 (71.6)	<0.001*
	Yes	3 (40.5)	9 (28.4)	
Dyslipidemia	No	545 (98.8)	298 (99.6)	<0.001*
	Yes	837 (1.2)	321 (0.4)	
Obesity	No	1263 (75.3)	577 (66.4)	0.165
	Yes	119 (24.7)	42 (33.6)	
Family History of Premature Death	No	1309 (28.5)	585 (48.8)	0.847
	Yes	73 (71.5)	34 (51.2)	
Chronic Kidney Disease	No	1023 (90.6)	517 (92.7)	<0.001*
	Yes	359 (9.4)	102 (7.3)	
History of ASCVD	No	180 (87.5)	89 (97.9)	0.324
	Yes	860 (12.5)	369 (2.1)	
Atrial Fibrillation	No	702 (89.4)	321 (70.5)	0.321
	Yes	338 (10.6)	137 (29.0)	
History of Implanted Device	No	999 (13.5)	439 (22.2)	0.851
	Yes	41 (86.5)	19 (77.8)	
History of Structural Heart Disease	No	988 (65.3)	426 (67.4)	0.124
	Yes	52 (34.7)	32 (32.6)	
History of Heart Failure	No	238 (95.7)	119 (96.1)	0.260
	Yes	1139 (4.3)	495 (3.9)	
Type of Heart Failure	HFpEF	329 (94.3)	920 (39.5)	0.006*
	HFrEF	20 (5.7)	64 (6.5)	
Mechanical Ventilation	No	889 (15.8)	342 (19.3)	0.624
	Yes	43 (84.2)	19 (80.7)	
Number of Hospital Admissions[a]	0	546 (39.1)	249 (41.1)	0.966
	1	268 (21.2)	118 (18.2)	
	2	106 (10.1)	45 (7.0)	
	>2	173 (29.6)	76 (33.6)	
Death	No	1249 (93.8)	565 (95.3)	0.365
	Yes	139 (6.2)	54 (4.7)	
Ejection Fraction (%), mean ± SD		39 ± 12.6	36.6 ± 13	<0.001*

(*Continued*)

Table 1 (Continued).

Variable, n (%)		DM (n=1388)	No DM (n=619)	p-value
PASP (mmHg), mean ± SD		26.7 ± 22.8	27.8 ± 22.8	0.570
Echo LVH (mm), mean ± SD		16.9 ± 27.3	15.8 ± 20	0.544
Echo LAE (mm)		25.1 ± 19.5	26.8 ± 20	0.125
SBP (mmHg)	Normal	756	391	0.016*
	>140	77	22	
DBP (mmHg)	Normal	951	441	0.388
	>90	64	24	
Length of Hospital Stay, mean ± SD		6.28 ± 7.42	6.22 ± 7.84	0.152

Notes: *Statistical significance was determined with a p-value ≤ 0.05.
Abbreviations: DM, Diabetes Mellitus; PND, Paroxysmal Nocturnal Dyspnea; BMI, Body Mass Index; ASCVD, Atherosclerotic Cardiovascular Disease; HFpEF, Heart Failure with preserved Ejection Fraction; HFrEF, Heart Failure with reduced Ejection Fraction; SD, Standard Deviation; PASP, Pulmonary Arterial Systolic Pressure; LVH, Left Ventricular Hypertrophy; LAE, Left Atrial Enlargement; SBP, Systolic Blood Pressure; DBP, Diastolic Blood Pressure.

on lipid profiles, with DM patients showing a higher incidence of abnormal cholesterol and LDL levels as detailed in Table 2 ($p = 0.008$ and $p = 0.003$, respectively). Hemoglobin levels were also notably different, with DM patients more frequently presenting with levels below the defined normal range ($p < 0.001$). Kidney function test results were significant for the Estimated Glomerular Filtration Rate (eGFR), showing more advanced stages of renal impairment in the DM group ($p = 0.006$).

Comparative Analysis by Diabetes Status and Ejection Fraction

When categorized by diabetes status and ejection fraction, the study identified significant differences in sex distribution, age, and prevalence of hypertension among the groups as presented in Table 3 ($p = 0.011$, $p = 0.002$, and $p = 0.020$, respectively). There was a notable difference in the prevalence of atrial fibrillation between the No DM with HFrEF group and the DM with HFrEF group ($p = 0.040$). However, no significant differences were observed in the history of implanted devices or structural heart disease across different categories. Additionally, mortality varied significantly between the DM with HFrEF and the No DM with HFrEF groups ($p = 0.044$).

Mortality Prediction Models

The performance metrics of the mortality prediction algorithms are summarized in Table 4. The Random Forest Classifier emerged as the top performer with the highest accuracy of 90.02% and an AUC of 80.51%, suggesting it is the most effective model for predicting mortality among heart failure patients (Figure 1). Despite its lower sensitivity at 32.56%, its specificity of 96.39% indicates excellent capability in correctly identifying patients who will not experience the event (mortality), making it highly reliable in negative predictions. The Logistic Regression model, known for its interpretability, also demonstrated good performance with a sensitivity of 72.09%, the highest among the models, which makes it particularly useful in identifying high-risk patients correctly. It has a moderate specificity of 73.97% and an AUC of 79.15%, indicating decent overall performance. The Support Vector Machine showed a balanced profile with an accuracy of 80.74% and a relatively lower AUC of 73.65% compared to other models. It has a sensitivity of 46.51% and a specificity of 84.54%, positioning it as a solid choice for predicting mortality with reasonable confidence in identifying true negatives. Lastly, the eXtreme Gradient Boosting model matched the Random Forest in accuracy at 90.02% and presented an AUC of 78.21%. It has a sensitivity of 39.53% and a very high specificity of 95.62%, like the Random Forest, highlighting its strength in specificity but lower performance in sensitivity.

Table 2 Laboratory Variables of Diabetes and Non-Diabetes Among Patients with Heart Failure

Laboratory Findings	DM (n=1388)	No DM (n=619)	p-value
Cholesterol			0.008*
Normal	460	182	
High	63	44	0.826
HDL			
Normal	140	63	
Low	379	164	
LDL			0.003*
Normal	475	192	
High	52	41	
Triglyceride			0.021*
Normal	316	158	
High	209	71	
BNP			0.298
Normal	30	16	
High	678	260	
NT-proBNP			0.176
Normal	14	11	
High	195	87	
NA			0.017*
<136	424	150	
136–145	850	400	
>145	43	26	
K			0.056
<3.5	68	35	
3.5–5	1058	483	
>5	188	60	
Hemoglobin			<0.001*
Normal	406	262	
<14 Male or Female <12	818	296	
Estimated GFR			0.006*
Stage 5	24	6	
Stage 4	53	17	
Stage 3b	74	13	
Stage 3a	47	25	
Stage 2	126	68	
Stage 1	97	50	
Blood Urea Nitrogen			0.243
Normal	282	135	
>20	912	379	
Creatinine			<0.001*
Normal	645	380	
High	625	175	

Notes: *Statistical significance was determined with a p-value ≤ 0.05.
Abbreviations: DM, Diabetes Mellitus; HDL, High Density Lipoprotein; LDL, Low Density Lipoprotein; BNP, B-type Natriuretic Peptide; N-terminal prohormone of brain natriuretic peptide; NA, Sodium levels; K, Potassium levels; GFR, Glomerular Filtration Rate.

Permutation Feature Importance

Permutation feature importance analysis as detailed in Figure 2, identified creatinine >115 µmol/L as the variable with the highest mean decrease in model performance when omitted, underscoring its importance in mortality prediction in

Table 3 Comparative Analysis of Clinical Characteristics and Outcomes Among Patients with Heart Failure Categorized by Diabetes Mellitus Status and Ejection Fraction

Variable	DM with HFrEF (815)	DM with HFpEF (386)	No DM with HFrEF (402)	No DM with HFpEF (138)	p-value
Sex					0.011*
Male	459	212	258	69	
Female	345	170	141	66	
Age					0.002*
<40	19	15	23	6	
40–49	47	24	38	5	
50–59	115	49	73	19	
60–69	209	100	93	26	
≥70	359	164	148	73	
BMI					0.050*
Normal	111	62	71	21	
Overweight	173	79	100	20	
Obesity	170	94	68	33	
Clinical Presentation					
Fatigue	507	235	266	76	0.057
Dyspnea	99	44	47	19	0.756
Orthopnea	429	197	241	68	0.004*
Chest pain	238	474	236	81	0.106
Palpitations	657	305	315	98	0.706
PND	231	498	261	81	0.074
Hypertension					0.020*
No	132	70	89	18	
Yes	636	293	279	107	
Smoking					0.363
No	538	244	239	86	
Yes	230	119	129	39	
Alcohol					0.144
No	764	361	365	122	
Yes	4	2	3	3	
Dyslipidemia					0.599
No	315	138	158	51	
Yes	453	225	210	74	
Obesity					0.598
No	705	326	339	116	
Yes	63	37	29	9	
Family History of Premature death					0.242
No	713	345	346	121	
Yes	55	18	22	4	
Chronic Kidney Disease					0.395
No	573	282	290	97	
Yes	195	81	78	28	
History of ASCVD					0.418
No	114	50	57	24	
Yes	507	236	242	72	
Atrial Fibrillation					0.040*
No	408	207	214	58	
Yes	213	79	85	38	

(*Continued*)

Table 3 (Continued).

Variable	DM with HFrEF (815)	DM with HFpEF (386)	No DM with HFrEF (402)	No DM with HFpEF (138)	p-value
History of Implanted Device					0.903
No	597	276	287	91	
Yes	24	10	12	5	
History of Structural Heart Disease					0.984
No	583	269	279	90	
Yes	38	17	20	6	
History of Heart Failure					0.747
No	145	71	64	27	
Yes	662	312	332	110	
Number of Hospital Admissions[a]					0.641
0	327	148	162	66	
1	153	74	87	23	
2	60	32	29	13	
>2	96	51	43	11	
Mechanical Ventilation					0.555
No	524	237	218	82	
Yes	20	13	13	5	
Death					0.044
No	718	355	372	127	
Yes	96	31	30	11	0.457
PASP (mm Hg), mean ± SD	27.8 ± 22	25.4 ± 23	27.8 ± 22	23.4 ± 23	
Echo LVH (mm), mean ± SD	15.7 ± 19	15.8 ± 19.7	14.6 ± 19	18.2 ± 26.3	0.719
Echo LAE (mm)	24.8 ± 19.8	24.4 ± 19.7	25.9 ± 19.9	24.2 ± 19.5	0.774
SBP (mmHg)					0.990
Normal	452	225	238	80	
>140	38	20	19	7	
DBP (mmHg)					0.637
Normal	560	274	279	93	
>90	31	18	21	4	
Length of Hospital Stay, mean ± SD	6.20 ± 7.62	6.93 ± 7.74	6 ± 7.43	6.92 ± 9.66	0.403

Notes: *Statistical significance was determined with a p-value ≤ 0.05.
Abbreviations: DM, Diabetes Mellitus; HFrEF, Heart Failure with reduced Ejection Fraction; HFpEF, Heart Failure with preserved Ejection Fraction; BMI, Body Mass Index; PND, Paroxysmal Nocturnal Dyspnea; ASCVD, Atherosclerotic Cardiovascular Disease; PASP, Pulmonary Arterial Systolic Pressure SD, Standard Deviation; LVH, Left Ventricular Hypertrophy; LAE, Left Atrial Enlargement; SBP, Systolic Blood Pressure; DBP, Diastolic Blood Pressure.

Table 4 Performance Metrics of Proposed Algorithms for Mortality Prediction

Model	Accuracy	AUC	Sensitivity	Specificity
Random Forest Classifier	90.02%	80.51%	32.56%	96.39%
Logistic Regression	73.78%	79.15%	72.09%	73.97%
Support Vector Machine	80.74%	73.65%	46.51%	84.54%
eXtreme Gradient Boosting	90.02%	78.21%	39.53%	95.62%

Abbreviation: AUC, Area Under the Curve.

heart failure patients. The length of hospital stays and requirement for mechanical ventilation were also identified as significant predictors, followed by chronic kidney disease, dyslipidemia, sex (male), and Blood Urea Nitrogen >20 mg/dL.

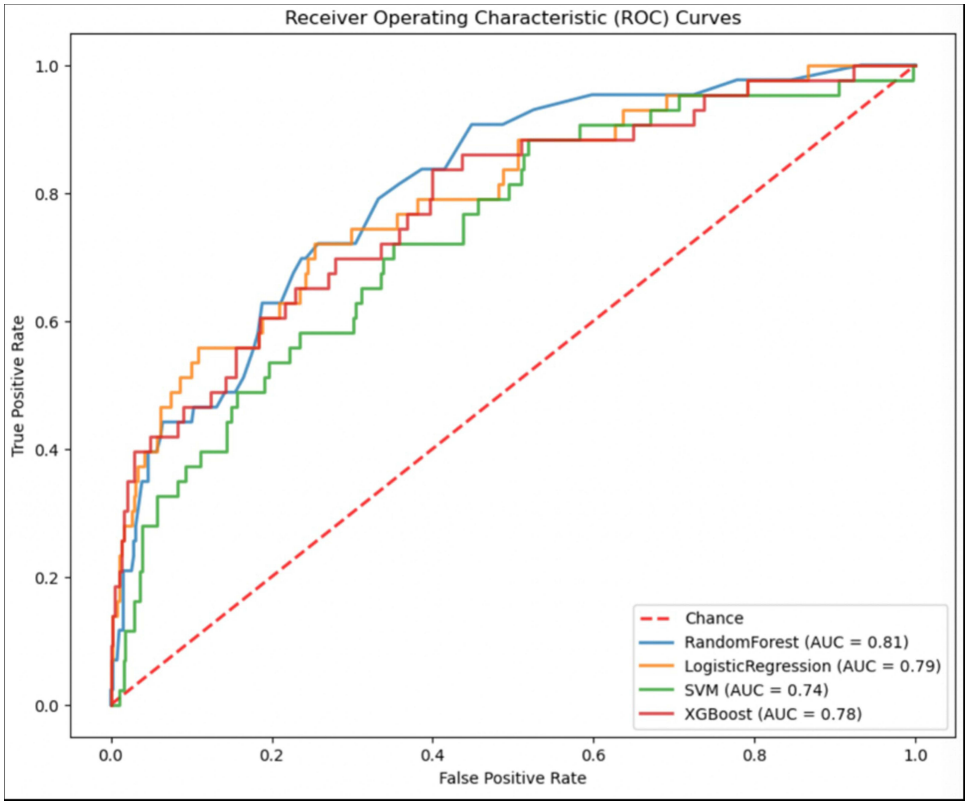

Figure 1 The dashed line represents the chance discrimination level. Each model's AUC was noted, with the Random Forest Classifier and Logistic Regression outperforming the SVM and XGBoost models marginally.

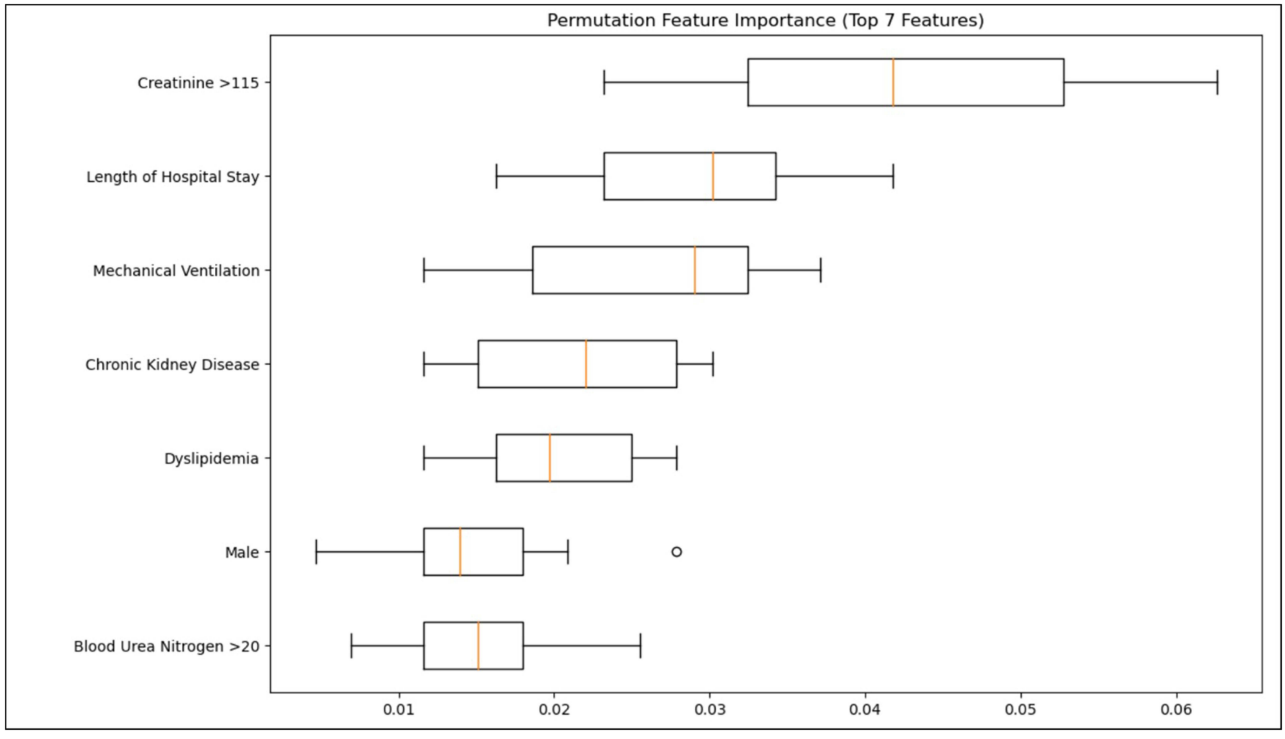

Figure 2 The error bars represent the standard deviation of the permutation importance over multiple shuffles.

Discussion

Our comprehensive analysis of 2007 heart failure (HF) patients, distinguishing between those with diabetes mellitus (DM) and those without DM, provides critical insights into the clinical and laboratory characteristics that define these cohorts. This detailed exploration reinforces the complex relationship between DM and HF. The prevalence of DM in our HF cohort underscores a notable intersection, aligning with the findings of previous studies that suggest a significant overlap between these conditions.[13,14] The demographic findings from our study—highlighting an increased incidence of HF among males, older individuals, and those classified as obese within the DM cohort mirror global trends and emphasize the multifaceted risk factors contributing to HF in the DM population.[15,16] These trends are critical for clinicians to consider, as they suggest that targeted screening and intervention strategies could significantly benefit these high-risk groups. Targeted screening involves systematically identifying individuals based on specific risk factors, such as diabetes, age, and obesity, that predispose them to heart failure. Subsequent interventions are then tailored to address these specific risks, potentially including more aggressive management of diabetes, lifestyle modifications, and closer monitoring for heart failure symptoms, to improve outcomes and prevent disease progression. Our findings on the higher prevalence of HF with preserved ejection fraction (HFpEF) among patients with DM contribute to an evolving narrative in the literature, suggesting distinct pathophysiological mechanisms in HF development among diabetic patients.[17,18] This is further supported by our laboratory findings, where abnormal cholesterol and LDL levels, alongside reduced eGFR, highlight the systemic impact of DM on cardiovascular health. These findings echo the significance of comprehensive cardiovascular risk management in patients with DM, as evidenced by multiple studies.[19,20] The mortality prediction models evaluated in our study, particularly the Random Forest Classifier, demonstrated high accuracy and area under the curve (AUC) metrics, with specificity and sensitivity analyses offering nuanced insights into their clinical utility. The logistic regression model's 72.09% sensitivity indicates its strength in correctly identifying patients at higher risk of death, a crucial capability for proactive patient management. The specificity of these models, though not explicitly stated, complements this sensitivity by identifying those not at risk and minimizing unnecessary interventions. The predictive power of these models, especially when considering variables like creatinine levels and the requirement for mechanical ventilation, is supported by the broader literature. Studies have consistently highlighted the role of kidney function markers in predicting outcomes in HF patients, with creatinine serving as a pivotal indicator of cardiovascular risk.[21,22] The inclusion of mechanical ventilation requirements reflects the severity of acute exacerbations, aligning with research suggesting its predictive value in HF mortality.[23] The specificity and sensitivity of our mortality prediction models have significant clinical implications. The ability to accurately predict mortality risk in HF patients, particularly those complicated by DM, can guide clinicians in prioritizing interventions for those at greatest risk. This approach not only enhances patient care but also optimizes resource allocation within healthcare systems. The high accuracy and AUC of the Random Forest Classifier suggest it may be best suited for integration into clinical decision-making processes, offering a robust tool for risk stratification and management planning. Further research is needed to refine these models, particularly by expanding the variables included in the analysis to encompass a wider range of clinical and social determinants of health. Future studies should consider integrating socio-economic factors such as education, income, and employment status, genetic markers linked to heart failure and diabetes mellitus, detailed dietary and lifestyle data, inflammatory markers like C-reactive protein, and data on medication adherence. These additions will enhance the models' capacity to account for the complex interactions between clinical outcomes and broader social and biological factors, potentially improving the predictive accuracy and clinical usefulness of our findings. Our study should be interpreted with caution in the context of several limitations. First, our analysis is based on single-country data from Jordan, which may limit the generalizability of the findings to populations with different demographic and health system characteristics. Secondly, the study utilizes registry data, which, while comprehensive, is prone to potential biases such as missing information and reporting inaccuracies. Furthermore, not all desirable variables were consistently recorded or available for all patients, possibly affecting the robustness of our conclusions. Additionally, the relatively short follow-up duration in some cases may not fully capture long-term outcomes and survival predictors. Another significant limitation is that while HbA1c values were used to confirm the presence of diabetes mellitus, we did not collect or analyze these values to assess the severity of diabetes across all centers, which precluded a detailed examination of the impact of

diabetes severity on heart failure outcomes. Lastly, while our models showed promising results within our dataset, their performance has not yet been validated externally, a crucial step to assess their real-world applicability and reliability. Future studies should aim to validate these models in diverse populations and settings, ensuring their applicability across different healthcare systems and patient demographics.

Conclusion

This study reveals notable differences between diabetic and non-diabetic heart failure patients in Jordan, showing poorer outcomes for those with diabetes, particularly higher prevalence of HFpEF and worse renal and lipid profiles. High creatinine levels, longer hospital stays, and mechanical ventilation needs were significant mortality predictors. The utility of our machine learning models, especially the Random Forest Classifier and Logistic Regression, varied. The Random Forest demonstrated high accuracy and specificity, ideal for minimizing false positives, while Logistic Regression, with higher sensitivity, proved better at identifying high-risk patients. These models, although promising in mortality prediction, should be complementarily used with clinical assessments to enhance decision-making in patient care. Further validation and refinement of these models are crucial for improving their accuracy and real-world applicability.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Ponikowski P, Voors AA, Anker SD. et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2016;37(27):2129–2200. doi:10.1093/eurheartj/ehw128
2. Martín-Timón I, Sevillano-Collantes C, Segura-Galindo A. Type 2 diabetes and cardiovascular disease: have all risk factors the same strength? *World J Diabetes*. 2014;5(4):444. doi:10.4239/wjd.v5.i4.444
3. Rodgers JL, Jones J, Bolleddu SI, et al. Cardiovascular risks associated with gender and aging. *J Cardiovasc Dev Dis*. 2019;6(2). doi:10.3390/jcdd6020019
4. Ma CX, Ma XN, Guan CH, Li YD, Mauricio D, Fu SB. Cardiovascular disease in type 2 diabetes mellitus: progress toward personalized management. *Cardiovasc Diabetol*. 2022;21(1). doi:10.1186/s12933-022-01516-6
5. Tousoulis D, Oikonomou E, Siasos G, Stefanadis C. Diabetes mellitus and heart failure. *Eur Cardiol*. 2014;9(1):37–42. doi:10.15420/ecr.2014.9.1.37
6. Johansson I, Edner M, Dahlström U, Näsman P, Rydén L, Norhammar A. Is the prognosis in patients with diabetes and heart failure a matter of unsatisfactory management? An observational study from the Swedish heart failure registry. *Eur J Heart Fail*. 2014;16(4):409–418. doi:10.1002/ejhf.44
7. Maddox TM, Januzzi JL, Allen LA, et al. 2021 update to the 2017 ACC expert consensus decision pathway for optimization of heart failure treatment: answers to 10 pivotal issues about heart failure with reduced ejection fraction: a report of the American College of Cardiology solution set oversight committee. *J Am Coll Cardiol*. 2021;77(6):772–810. doi:10.1016/j.jacc.2020.11.022
8. Lamari M, Azizi N, Hammami NE, et al. SMOTE–ENN-based data sampling and improved dynamic ensemble selection for imbalanced medical data classification. *Adv Intell Syst Comput*. 2021;1188:37–49.
9. Svetnik V, Liaw A, Tong C, Christopher Culberson J, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci*. 2003;43(6):1947–1958. doi:10.1021/ci034160g
10. Nettleton D. Data Modeling. *Commercial Data Mining*. 2014;2:137–157.
11. Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006;24(12):1565–1567. doi:10.1038/nbt1206-1565
12. Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorobot*. 2013;7:63623. doi:10.3389/fnbot.2013.00021
13. Van Den Berge JC, Constantinescu AA, Boiten HJ, Van Domburg RT, Deckers JW, Akkerhuis KM. Short- and long-term prognosis of patients with acute heart failure with and without diabetes: changes over the last three decades. *Diabetes Care*. 2018;41(1):143–149. doi:10.2337/dc17-0544
14. Tromp J, Tay WT, Ouwerkerk W, et al. Multimorbidity in patients with heart failure from 11 Asian regions: a prospective cohort study using the ASIAN-HF registry. *PLoS Med*. 2018;15:3.
15. Shah R, Gayat E, Januzzi JL, et al. Body mass index and mortality in acutely decompensated heart failure across the world: a global obesity paradox. *J Am Coll Cardiol*. 2014;63(8):778–785. doi:10.1016/j.jacc.2013.09.072
16. Dewan P, Jhund PS, Shen L, et al. Heart failure with reduced ejection fraction: comparison of patient characteristics and clinical outcomes within Asia and between Asia, Europe and the Americas. *Eur J Heart Fail*. 2019;21(5):577–587. doi:10.1002/ejhf.1347
17. Dodani S, Sharma GK. Presence of coronary artery disease in diabetic and non diabetic South Asian immigrants. *Indian Heart J*. 2018;70(1):50–55. doi:10.1016/j.ihj.2017.07.009
18. Bank IEM, Gijsberts CM, Teng THK, et al. Prevalence and clinical significance of diabetes in asian versus white patients with heart failure. *JACC Heart Fail*. 2017;5(1):14–24. doi:10.1016/j.jchf.2016.09.015
19. Löfman I, Szummer K, Hagerman I, Dahlström U, Lund LH, Jernberg T. Prevalence and prognostic impact of kidney disease on heart failure patients. *Open Heart*. 2016;3(1):e000324. doi:10.1136/openhrt-2015-000324

20. Goh VJ, Tromp J, Teng THK, et al. Prevalence, clinical correlates, and outcomes of anaemia in multi-ethnic Asian patients with heart failure with reduced ejection fraction. *ESC Heart Fail*. 2018;5(4):570. doi:10.1002/ehf2.12279
21. Harikrishnan S, Sanjay G, Anees T, et al. Clinical presentation, management, in-hospital and 90-day outcomes of heart failure patients in trivandrum, Kerala, India: the trivandrum heart failure registry. *Eur J Heart Fail*. 2015;17(8):794–800. doi:10.1002/ejhf.283
22. Iribarren C, Karter AJ, Go AS, et al. Glycemic control and heart failure among adult patients with diabetes. *Circulation*. 2001;103(22):2668–2673. doi:10.1161/01.CIR.103.22.2668
23. Biswas A, Singh SK, Singh RK. Linkages between hypertension and coronary heart disease in India: evidence from India human development survey-2 (2011–2012). *Indian J Community Med*. 2017;42(4):200. doi:10.4103/ijcm.IJCM_168_16