IVNA GIRARD CUNHA VIEIRA LIMA

Prognostic outcomes in chronic heart failure: a retrospective population-based cohort study from a Brazilian tertiary center

Desfechos da insuficiência cardíaca crônica: análise baseada na população de uma instituição brasileira: Estudo de coorte retrospectivo em pacientes com insuficiência cardíaca crônica para investigar os desfechos e explorar fatores de risco e prognóstico associados

São Paulo

2026

FACULDADE DE MEDICINA

UNIVERSIDE DE SÃO PAULO

IVNA GIRARD CUNHA VIEIRA LIMA

Prognostic outcomes in chronic heart failure: a retrospective population-based cohort study from a Brazilian tertiary center

Desfechos da insuficiência cardíaca crônica: análise baseada na população de uma instituição brasileira: Estudo de coorte retrospectivo em pacientes com insuficiência cardíaca crônica para investigar os desfechos e explorar fatores de risco e prognóstico associados

Tese apresentada à Faculdade de Medicina da Universidade de São Paulo para obtenção de título de Doutor em Ciências: Programa de Cardiologia

Orientador: Prof. Dr. Edimar Alcides Bocchi

São Paulo
2026

CATALOGING IN PUBLICATION DATA

Faculdade de Medicina da Universidade de São Paulo Central Library

Lima, Ivna Girard Cunha Vieira

Prognostic outcomes in chronic heart failure: a retrospective population-based cohort study from a Brazilian tertiary center / Ivna Girard Cunha Vieira Lima; Edimar Alcides Bocchi, advisor. -- São Paulo, 2026.

Thesis (Doctor of Science Degree) – Cardiology Program. Faculdade de Medicina da Universidade de São Paulo, 2026.

1.Heart failure. 2.Reduced ejection fraction 3.Prognosis 4.Biomarkers 5.Cohort studies I. Bocchi, Edimar Alcides, adivsor II. Title

USP/FM/DBD - 048/26

Responsible: Daniela Amaral Barbosa, CRB-8 7533

To my mother, Veranilde Pereira Cunha

To my father, Atalicio Vieira Lima

To my brother, Ighor Cunha Vieira Lima

ACKNOWLEDGMENTS

This doctoral journey has been more than a scientific pursuit; it has been a profound lesson in resilience, collaboration, and the transformative power of shared purpose. The support I received not only shaped the scholar within but also influenced the person I have become — reminding me that every accomplishment carries the silent imprint of many hands and hearts.

My deepest appreciation goes to **Professor Edimar Alcides Bocchi**, whose mentorship, intellectual rigor, and inspiring dedication to heart failure research have profoundly guided my path. His example of scientific excellence and human integrity has been a true compass throughout this endeavor.

To the **Heart Failure Team at the Instituto do Coração (InCor)** of the Hospital das Clínicas, Faculty of Medicine, University of São Paulo, my heartfelt thanks for creating an environment where science, teaching, and care intertwine in the pursuit of better lives. The collaborative spirit at InCor has provided fertile ground for curiosity, growth, and discovery.

To my **colleagues, friends, and research collaborators**, I am deeply grateful for the meaningful exchanges, the endless discussions, and the shared commitment to advancing knowledge. Your insights and generosity of spirit have enriched this work in ways that transcend data and manuscripts.

To my **family and loved ones**, whose love, patience, and unwavering belief sustained me through uncertainty and exhaustion — your support has been my anchor and the foundation upon which all of this stands. Every step of this path carries a reflection of your strength.

“Eu sou mais forte que eu”.

Clarisse Lispector

RESUMO

Lima IGCV. Desfechos da insuficiência cardíaca crônica: análise baseada na população de uma instituição brasileira: Estudo de coorte retrospectivo em pacientes com insuficiência cardíaca crônica para investigar os desfechos e explorar fatores de risco e prognóstico associados [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2026.

A insuficiência cardíaca com fração de ejeção reduzida (ICFER) é uma síndrome clínica complexa e biologicamente heterogênea, na qual o prognóstico e a resposta terapêutica resultam da interação entre remodelamento estrutural, carga de comorbidades, dinâmica dos biomarcadores e tolerância ao tratamento. Estratégias tradicionais de estratificação de risco, baseadas exclusivamente na fração de ejeção do ventrículo esquerdo e na adesão às diretrizes, não capturam plenamente essa complexidade, especialmente em populações do mundo real de países de baixa e média renda. O objetivo desta tese foi investigar determinantes clínicos, laboratoriais, de imagem e terapêuticos associados aos desfechos em pacientes ambulatoriais com insuficiência cardíaca crônica acompanhados em um centro terciário brasileiro, com ênfase em fenótipos clínicos, distúrbios eletrolíticos e potencial de remodelamento reverso. Trata-se de um estudo de coorte retrospectivo, de base populacional, que incluiu mais de 10.000 pacientes com ICFER acompanhados em ambulatório especializado. Foram realizadas análises multidimensionais que contemplaram: (1) a associação entre a exposição à terapia farmacológica modificadora da doença e o risco de hipercalemia e de mortalidade; (2) a identificação de fenótipos clínicos por aprendizado de máquina não supervisionado; (3) a caracterização estrutural por ressonância magnética cardíaca, com o desenvolvimento de um escore contínuo de probabilidade de remodelamento reverso; e (4) a avaliação do desempenho prognóstico do peptídeo natriurético tipo B (BNP) em subgrupos clínicos relevantes. Foram identificados cinco fenótipos distintos de ICFER, cada um com perfis específicos de comorbidades, de uso de terapias e de trajetórias de mortalidade e hospitalização. Os fenótipos chagásico e cardiorrenal apresentaram risco desproporcionalmente elevado de desfechos adversos, independentemente da fração de ejeção. A intensificação dos inibidores do sistema renina–angiotensina associou-se a maior risco de hipercalemia, enquanto os betabloqueadores demonstraram benefício de sobrevida dependente do tempo. O escore de remodelamento reverso, derivado da ressonância magnética cardíaca, mostrou forte associação com mortalidade e superou a fração de ejeção na estratificação prognóstica. O BNP manteve valor prognóstico consistente, porém com limiares de risco dependentes do contexto clínico, variando conforme obesidade, doença renal crônica e fibrilação atrial. Conclui-se que o prognóstico na ICFER é determinado por uma interação complexa entre o fenótipo clínico, a capacidade estrutural de recuperação, os biomarcadores e a tolerabilidade terapêutica. Uma abordagem multidimensional e orientada por fenótipos pode aprimorar a estratificação de risco, otimizar a tomada de decisão clínica e contribuir para a implementação de estratégias de medicina de precisão em populações de pacientes com insuficiência cardíaca heterogêneas.

Palavras-chave: Insuficiência cardíaca., Fração de ejeção reduzida, Prognóstico, Biomarcadores, Estudos de coorte

ABSTRACT

Lima IGCV. Prognostic outcomes in chronic heart failure: a retrospective population-based cohort study from a Brazilian tertiary center [thesis]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2026.

Background: Heart failure with reduced ejection fraction (HFrEF) is a biologically heterogeneous syndrome in which prognosis and therapeutic responseare shaped by structural remodeling, comorbidity burden, biomarker dynamics, and treatment tolerability. Traditional risk stratification based on left ventricular ejection fraction (LVEF) and guideline adherence fails to fully capture this complexity, particularly in real-world populations from low- and middle-income countries. **Objectives**: To investigate clinical, laboratory, imaging, and therapeutic determinants of prognosis in ambulatory patients with chronic HFrEF followed at a Brazilian tertiary center, with emphasis on phenotype-specific risk, electrolyte disturbances, and reverse remodeling potential. **Methods:** This retrospective population-based cohort study included more than 10,000 ambulatory HFrEF patients followed at a specialized heart failure clinic. Multidimensional analyses were performed across complementary domains: (1) guideline-directed medical therapy (GDMT) exposure and its association with hyperkalemia and mortality using inverse probability weighting and time-varying survival models; (2) unsupervised machine-learning phenotyping based on clinical and comorbidity profiles; (3) imaging-derived structural phenotypes and a continuous cardiac magnetic resonance–based reverse remodeling probability score; and (4) prognostic assessment of B type natriuretic peptide (BNP) across clinically relevant subgroups. **Results:** Five distinct HFrEF phenotypes were identified, each exhibiting divergent trajectories of mortality, hospitalization, and treatment utilization. Chagasic and cardiorenal phenotypes demonstrated disproportionately high risk despite differences in age and LVEF. Uptitration of renin- angiotensin system inhibitors was independently associated with increased hyperkalemia risk, while beta-blockers conferred time-dependent survival benefit. A novel imaging-derived reverse remodeling probability score robustly predicted mortality and outperformed LVEF.BNP remained a strong prognostic marker, although optimal risk thresholds varied substantially according to obesity, chronic kidney disease, and atrial fibrillation. **Conclusions:** Prognosis in chronic HFrEF is driven by a complex interplay of phenotype, structural remodeling capacity, biomarker context, and treatment tolerability. A phenotype-aware, multidimensional approach to risk stratification and therapy optimization may improve clinical decision-making and outcomes, particularly in heterogeneous real-world populations.

Keywords: Heart failure., Reduced ejection fraction, Prognosis, Biomarkers ,Cohort Studies

FIGURES INDEX

LIST OF ABBREVIATIONS

ACEI – Angiotensin-converting enzyme inhibitor
AF – Atrial fibrillation
ARB – Angiotensin receptor blocker
ARNI – Angiotensin receptor–neprilysin inhibitor
AUC – Area under the curve
BB – Beta-blocker
BMI – Body mass index
BNP – B-type natriuretic peptide
CI – Confidence interval
CKD – Chronic kidney disease
CMR – Cardiac magnetic resonance
COPD – Chronic obstructive pulmonary disease
CRP – C-reactive protein
CRS – Cardiorenal syndrome
eGFR – Estimated glomerular filtration rate
GDMT – Guideline-directed medical therapy
HF – Heart failure
HFpEF – Heart failure with preserved ejection fraction
HFmrEF – Heart failure with mildly reduced ejection fraction
HFrEF – Heart failure with reduced ejection fraction
HR – Hazard ratio
ICFER – Insuficiência cardíaca com fração de ejeção reduzida
InCor – Instituto do Coração
IPTW – Inverse probability of treatment weighting
IQR – Interquartile range
KM – Kaplan–Meier
LGE – Late gadolinium enhancement
LV – Left ventricle
LVEF – Left ventricular ejection fraction
LVEDV – Left ventricular end-diastolic volume
LVESV – Left ventricular end-systolic volume
MICE – Multivariate imputation by chained equations
MRA – Mineralocorticoid receptor antagonist
NLP – Natural language processing
NT-proBNP – N-terminal pro–B-type natriuretic peptide
NYHA – New York Heart Association
OS – Overall survival
PCA – Principal component analysis
RAAS – Renin–angiotensin–aldosterone system
RR – Reverse remodeling
RRp – Reverse remodeling probability score
RV – Right ventricle
SBP – Systolic blood pressure

SUMMARY

1 INTRODUCTION

1.1 PREFACE

This Ph.D. project's research was initiated in the Heart Failure Departament of the Instituto do Coração (InCor) at the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP) in São Paulo, Brazil, under the guidance of Professor Edimar Alcides Bocchi. Additionally, the student works with grants from Astra-Zeneca Brasil. Two scientific articles (chapters 3 to 5) resulting from this work have been published in peer-reviewed cardiovascular journals to this date.

1.2 HEART FAILURE

Heart failure (HF) is a systemic clinical syndrome in which the heart is unable to deliver sufficient blood flow to meet the metabolic requirements of peripheral tissues unless this is achieved at the cost of abnormally elevated intracardiac filling pressures. Rather than a single disease entity, HF represents the final expression of multiple cardiovascular pathologies, including ischemic heart disease, cardiomyopathies, valvular disorders, and hypertensive heart disease. Despite major therapeutic advances, HF continues to impose a major burden worldwide, driven by population aging and the increasing survival of patients after acute cardiovascular events (1,2,3).

Current epidemiologic estimates indicate that more than 64 million individuals are living with HF globally, with a rising trajectory in both prevalence and healthcare utilization (1,2). In parallel, HF remains one of the leading causes of hospitalization and cardiovascular mortality, creating a substantial economic and societal impact (1,2).

In Brazil, the epidemiological profile of HF broadly mirrors global trends but exhibits important regional and etiologic particularities. Data from large population-based cohorts such as ELSA-Brasil demonstrate that approximately 9% of Brazilian adults already exhibit symptomatic HF, while nearly half of the population falls into preclinical stages characterized

by structural heart disease or major risk factors (4). The high prevalence of hypertension, diabetes, obesity, and endemic Chagas disease contributes to this large reservoir of individuals at risk for HF progression (4).

Longitudinal national analyses suggest that although overall HF-related mortality and hospital admissions have declined over the past two decades, in-hospital case fatality rates and length of stay have paradoxically increased. This pattern likely reflects a shift toward hospitalization of patients with more advanced disease, delayed presentation, and substantial regional disparities in access to specialized cardiovascular care (5,6).

Figure 1 - Prevalence of heart failure (HF) stages by sex in Brazil

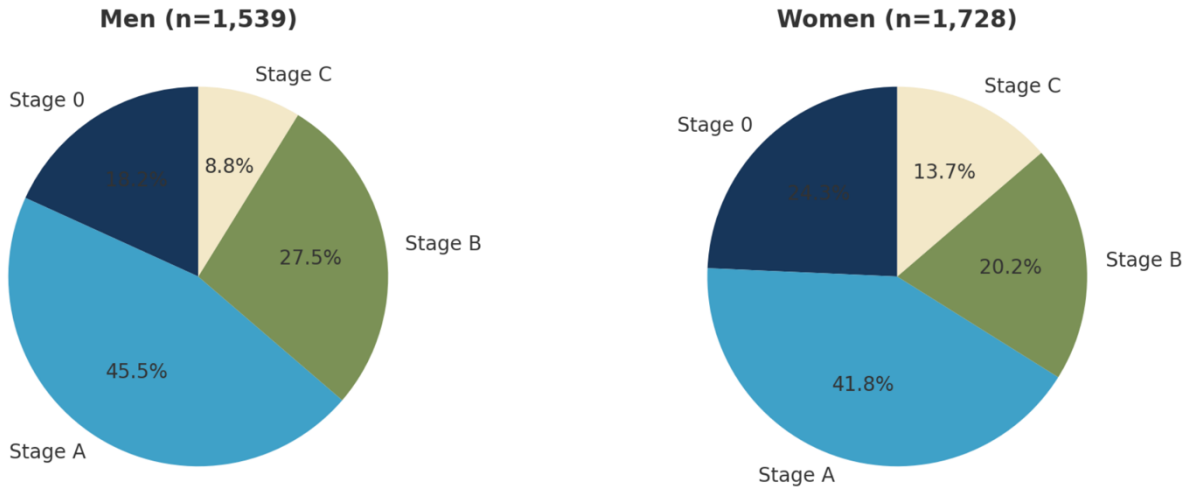


Figure1. Distribution of heart failure (HF) stages (Stage 0, A, B, and C) among men (n=1,539) and women (n=1,728) in the ELSA-Brasil cohort. Higher prevalence of Stages A and B was observed in both sexes. *Adapted from Pianca EG, Heidemann AI, Santos ÂBS, et al. Eur Heart J. 2024;45(Suppl_1).* https://doi.org/10.1093/eurheartj/ehae666.892

HF rarely occurs in isolation. Instead, it is typically accompanied by multiple cardiovascular and non-cardiovascular comorbidities that exert powerful effects on disease progression, treatment tolerance, and prognosis. Conditions such as diabetes mellitus, chronic kidney disease, chronic obstructive pulmonary disease, obesity, anemia, malignancy, and

chronic liver disease frequently coexist with HF and interact with its underlying pathophysiology (7).

These comorbidities amplify HF severity through overlapping biological pathways, including systemic inflammation, neurohormonal activation, oxidative stress, endothelial dysfunction, and metabolic derangement. Beyond direct myocardial effects, comorbid diseases often limit the use of life-saving therapies. For example, renal dysfunction may constrain renin–angiotensin–aldosterone system (RAAS) inhibition due to hyperkalemia, while pulmonary disease and frailty reduce exercise tolerance and rehabilitation capacity. Polypharmacy, altered drug metabolism, and competing risks further complicate long-term management (7).

Large registry studies reinforce the prognostic weight of multimorbidity. Analyses from the Swedish Heart Failure Registry, encompassing more than 30,000 patients, demonstrated that non-cardiac comorbidities significantly increased mortality in both HFrEF and HFpEF, although the relative impact differed by phenotype. Renal dysfunction, diabetes, and liver disease were particularly detrimental in HFrEF, whereas pulmonary disease exerted a stronger effect in HFpEF (8). Similarly, a nationwide Danish cohort of over 280,000 patients showed that individuals with multiple pre-existing comorbidities experienced disproportionately high early mortality, especially among younger patients, highlighting that biological age and disease burden are more relevant than chronological age alone (9).

From a clinical standpoint, multimorbidity undermines the implementation of guideline-directed medical therapy (GDMT) by increasing contraindications, adverse drug reactions, and treatment discontinuation. Cognitive impairment, depression, and social vulnerability further reduce adherence and self-care, contributing to recurrent decompensation and readmissions (7).

1.3 PHENOTYPES IN HEART FAILURE

The concept of phenotyping has emerged as a critical framework for understanding the heterogeneity of HF, particularly in heart failure with preserved ejection fraction (HFpEF). HFpEF encompasses multiple biological substrates, including metabolic dysfunction, atrial fibrillation, pulmonary vascular disease, and infiltrative cardiomyopathies. This diversity has driven the development of data-driven phenomapping strategies using cluster analysis and machine learning to identify patient subgroups with shared biological mechanisms and prognostic trajectories (10,11).

Figure 2 - Comorbidities and Pathophysiological Drivers in Heart Failure

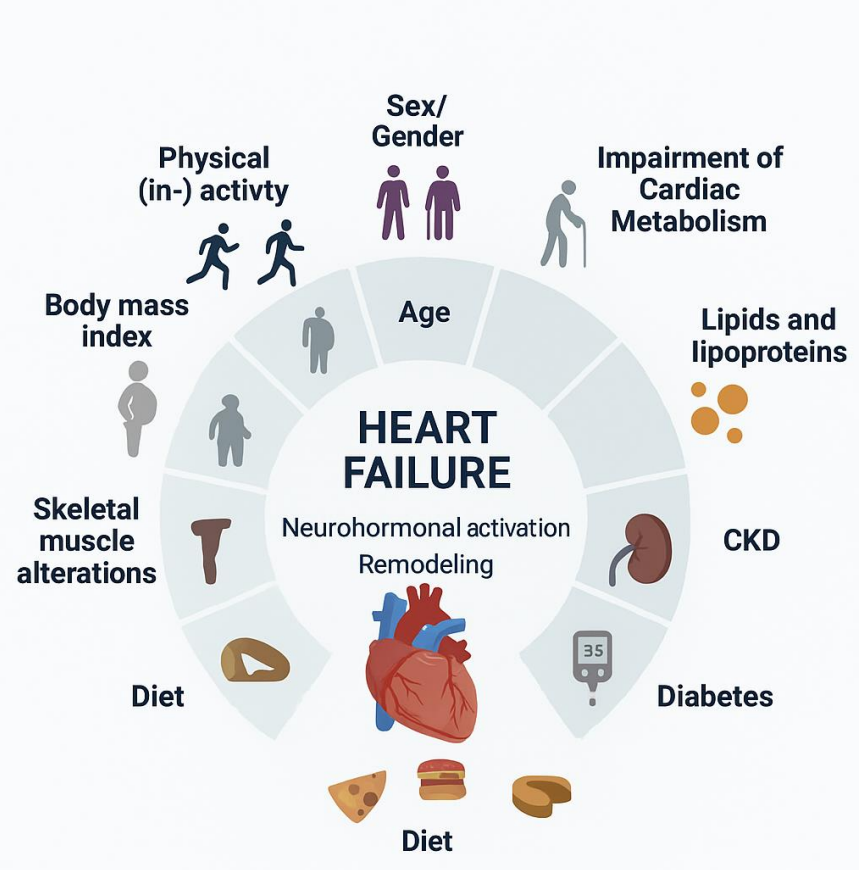


Figure2. Common clinical and metabolic conditions contributing to the development and progression of heart failure. Neurohormonal activation and remodeling are central mechanisms in the pathophysiological cascade. *Adapted from: Kintscher U, Edelmann F. The non-steroidal mineralocorticoid receptor antagonist finerenone and heart failure with preserved ejection fraction. Cardiovasc Diabetol. 2023;22. doi:10.1186/s12933-023-01899-0*

In contrast, heart failure with reduced ejection fraction (HFrEF) has traditionally been viewed as a relatively homogeneous syndrome defined primarily by systolic dysfunction, most often due to ischemic or dilated cardiomyopathy. As a result, treatment paradigms and risk stratification approaches have been applied uniformly across this population. However, this oversimplification fails to account for substantial intragroup variability in comorbidity burden, inflammatory activation, arrhythmic risk, renal dysfunction, and therapeutic responsiveness. Recent studies have challenged this paradigm by demonstrating that HFrEF patients can be segregated into distinct clinical and biological phenotypes with markedly different outcomes. Cluster analyses across the HF spectrum have identified subgroups characterized by congestion, metabolic disease, renal impairment, or inflammatory activation, each associated with unique prognostic profiles (13).

Advanced cardiometabolic and imaging-based phenotyping further underscores these differences. In HFpEF, markers such as glycated hemoglobin are more closely related to systemic inflammation and pulmonary congestion, whereas in HFrEF they correlate more strongly with sympathetic activation and hemodynamic compromise (14). Similarly, cardiovascular magnetic resonance has shown that patients with mid-range ejection fraction share structural and fibrotic features with both preserved and reduced phenotypes, blurring rigid LVEF-based categories (15).

Collectively, these observations indicate that ejection fraction alone does not capture the biological diversity of HF. A phenotype-oriented approach that integrates comorbidities, biomarkers, and structural remodeling is essential for advancing precision medicine across the HF continuum.

1.4 CARDIORENAL SYNDROME IN HEART FAILURE

Cardiorenal syndrome (CRS) describes a pathophysiologic continuum in which dysfunction of the heart and kidneys interact bidirectionally, such that acute or chronic impairment in one organ precipitates or worsens injury in the other. In the setting of heart failure (HF), renal dysfunction is highly prevalent and represents one of the strongest independent predictors of adverse outcomes, including hospitalization and mortality. The mechanisms underlying CRS extend well beyond reduced cardiac output and include venous congestion, neurohormonal activation, endothelial dysfunction, oxidative stress, and systemic inflammation, all of which contribute to progressive renal and myocardial injury (16).

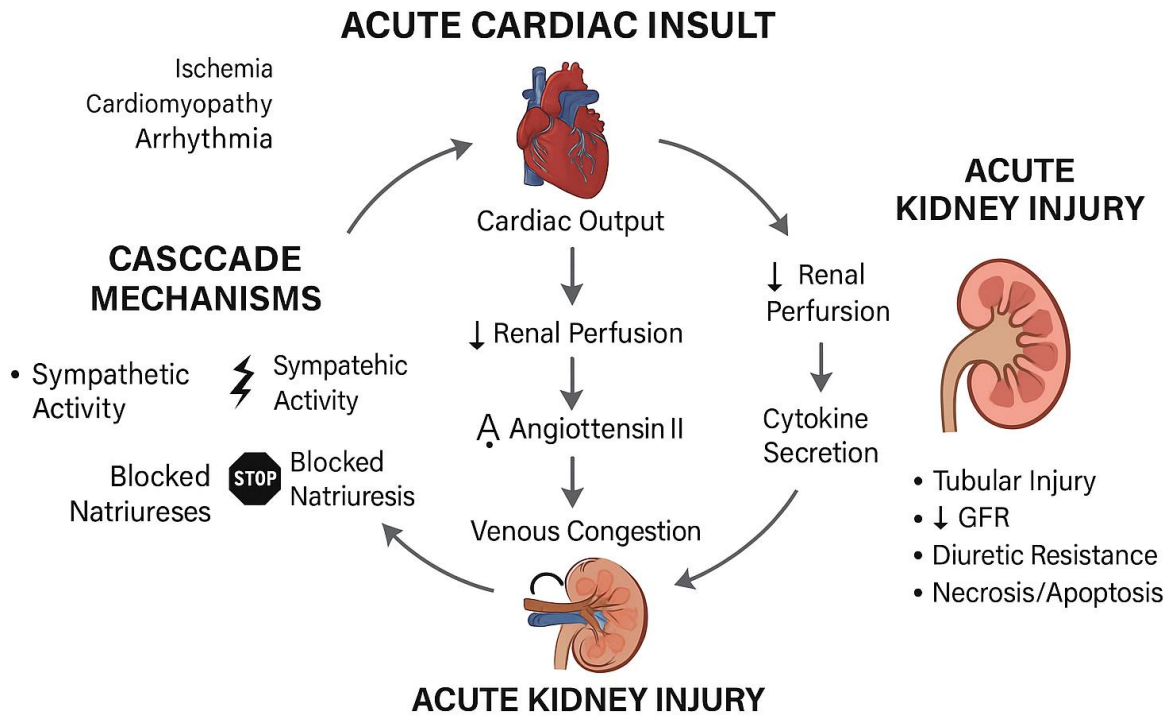
CRS is currently classified into five subtypes based on the initiating organ and the acute or chronic nature of dysfunction. In patients with HF, both acute and chronic forms of CRS are frequently encountered, often coexisting within the same individual. Nearly half of patients with HF exhibit some degree of renal impairment, and this subgroup experiences substantially higher rates of rehospitalization and death (17). Congestion-induced elevations in renal venous pressure, rather than reductions in arterial perfusion alone, have emerged as key determinants of declining kidney function in HF, reinforcing the central role of volume overload in CRS pathophysiology.

Importantly, transient worsening of renal function during aggressive diuresis does not necessarily portend worse prognosis when accompanied by effective decongestion. Clinical response and relief of congestion are more strongly associated with survival than short-term increases in serum creatinine, emphasizing the need for integrated interpretation of renal biomarkers in the context of volume status (16,17).

The development of CRS poses major therapeutic challenges, as renal dysfunction limits the safe use of disease-modifying therapies, particularly renin–angiotensin–aldosterone system

inhibitors and mineralocorticoid receptor antagonists. As a result, CRS represents both a marker of disease severity and a mediator of therapeutic underutilization.

Figure 3 - Pathophysiological Mechanisms Linking Acute Cardiac Insult to Acute Kidney Injury



Electrolyte abnormalities constitute a central manifestation of CRS in HF and significantly contribute to arrhythmic risk, therapeutic limitation, and adverse prognosis. Disturbances in sodium, potassium, and magnesium arise from impaired renal handling, neurohormonal activation, and the effects of commonly used HF medications.

Reduced effective circulating volume and renal hypoperfusion stimulate activation of the renin–angiotensin–aldosterone and sympathetic nervous systems, promoting sodium and water retention and leading to dilutional hyponatremia, a well-established marker of advanced HF and poor prognosis (16). Simultaneously, aldosterone-mediated potassium excretion and diuretic therapy predispose patients to hypokalemia, which increases the risk of ventricular arrhythmias and limits safe titration of neurohormonal blockade.

As renal function declines further, however, the kidney's ability to excrete potassium becomes impaired, particularly in the setting of RAAS inhibition or potassium-sparing diuretics, resulting in hyperkalemia. Hyperkalemia frequently leads to dose reduction or discontinuation of life-prolonging therapies, thereby indirectly worsening HF outcomes (17).

Magnesium deficiency, often underrecognized, commonly accompanies CRS due to renal losses and diuretic use and further increases susceptibility to arrhythmias and neuromuscular dysfunction. The dynamic interplay between renal impairment, pharmacologic therapy, and neurohormonal activation creates a narrow therapeutic window in which electrolyte homeostasis is fragile and requires continuous monitoring and individualized adjustment (18).

1.5 BIOMARKERS IN HEART FAILURE

Biomarkers play an essential role in the diagnosis, risk stratification, and therapeutic guidance of heart failure. They provide a window into the underlying pathophysiologic processes, including hemodynamic stress, neurohormonal activation, inflammation, and end-organ injury. Among the large array of investigated molecules, *B-type natriuretic peptides* (BNP and NT-proBNP) and *C-reactive protein* (CRP) have emerged as key biomarkers reflecting distinct but complementary biological pathways that drive disease progression and prognosis in HF.

1.5.1 INFLAMMATION AND HIGH-SENSITIVITY C-REACTIVE PROTEIN

Chronic low-grade inflammation is a hallmark of HF across etiologies and contributes directly to myocardial remodeling, endothelial dysfunction, and progressive ventricular impairment. Pro-inflammatory cytokines released by the failing myocardium stimulate hepatic production of CRP, making hs-CRP a sensitive indicator of systemic inflammatory activation in both stable and decompensated HF (19–21).

Elevated hs-CRP levels have been associated with the development of incident HF in population-based cohorts and with worse outcomes in established disease, including higher mortality and rehospitalization rates (19–21). Beyond its role as a marker, CRP may participate directly in myocardial injury by modulating complement activation, nitric oxide signaling, and microvascular function, thereby linking systemic inflammation to myocardial and vascular dysfunction.

Inflammatory activation is particularly pronounced in HF phenotypes characterized by obesity, diabetes, and chronic kidney disease, where metabolic and endothelial derangements perpetuate cytokine release and oxidative stress, further accelerating disease progression.

1.5.2 NATRIURETIC PEPTIDES AND NEUROHORMONAL STRESS

BNP and NT-proBNP reflect ventricular wall stress and neurohormonal activation, serving as endogenous counter-regulatory hormones that oppose sodium retention, vasoconstriction, and fibrosis. Released in response to myocardial stretch, these peptides promote natriuresis, vasodilation, and suppression of the RAAS and sympathetic nervous system, thereby mitigating hemodynamic overload (22,23).

Clinically, natriuretic peptides are integral to HF diagnosis, risk stratification, and therapeutic monitoring. Plasma concentrations correlate closely with intracardiac filling pressures and ventricular dysfunction and are incorporated into diagnostic algorithms for both acute and chronic HF (24,25). Serial measurements provide valuable prognostic information and may guide optimization of guideline-directed medical therapy (26,27).

Across HF phenotypes and care settings, elevated BNP and NT-proBNP consistently predict increased risk of mortality and rehospitalization, reinforcing their role as cornerstone biomarkers in contemporary HF management.

1.5.3 INTEGRATIVE AND PROGNOSTIC PERSPECTIVES

While natriuretic peptides primarily capture hemodynamic and neurohormonal stress, hs-CRP reflects systemic inflammatory burden. These two biological axes interact to shape disease trajectory, as inflammation may impair natriuretic peptide signaling and exacerbate vascular stiffness, while congestion and tissue hypoxia further amplify inflammatory pathways. Multimarker strategies that integrate BNP or NT-proBNP with hs-CRP improve prognostic discrimination beyond either biomarker alone (28,29).

This integrative framework supports a shift from a purely hemodynamic model of HF toward a multidimensional paradigm encompassing inflammatory, metabolic, renal, and neurohormonal domains. Such biomarker-guided phenotyping offers a foundation for precision medicine, enabling identification of patients most likely to benefit from targeted interventions and more intensive monitoring.

Figure 4 - Biomarkers in heart failure

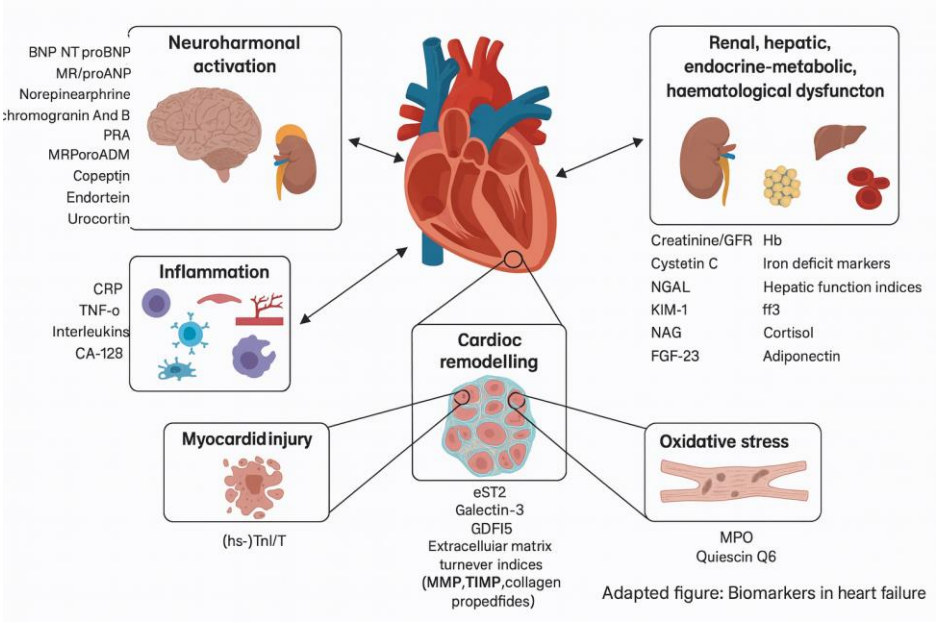


Figure 4. Schematic representation of key pathophysiological pathways and circulating biomarkers involved in heart failure. Central cardiac remodeling is influenced by neurohormonal activation, inflammation, myocardial injury, oxidative stress, and multi-organ dysfunction (renal, hepatic, endocrine-metabolic, and hematological). Each pathway is associated with representative biomarkers commonly used in clinical and research settings. *Adapted from the Castiglione, V., Aimo, A., Vergaro, G.* et al. *Biomarkers for the diagnosis and management of heart failure.* Heart Fail Rev *27, 625–643 (2022). https://doi.org/10.1007/s10741-021-10105-w*

1.6 HYPOTHESIS

1.6.1 General hypothesis

Among ambulatory patients with chronic heart failure (CHF) followed at a tertiary academic center in Brazil (InCor–HCFMUSP), clinical, laboratory, and imaging characteristics—together with biochemical markers of renal and electrolyte imbalance—are independently associated with mortality and hospitalization. Furthermore, adherence to guideline-directed medical therapy (GDMT) significantly influences long-term outcomes, while disturbances in potassium and sodium homeostasis reflect a distinct biological phenotype of risk.

1.6.2 Specific hypotheses

- Elevated serum potassium levels, even within the high-normal range, are independently associated with increased all-cause mortality and heart failure–related hospitalizations among ambulatory CHF patients.
- The cumulative incidence and recurrence of hyperkalemia follow a graded relationship with the intensity of neurohormonal blockade (particularly RAAS inhibitors, MRAs, and ARNi), and this relationship may differ across etiologic subgroups of heart failure.
- Progressive elevations in urea and creatinine, as well as a decline in estimated glomerular filtration rate (eGFR), predict adverse outcomes and reflect a cardiorenal interaction mediated by congestion and pharmacologic therapy.
- Reverse myocardial remodeling, identified through serial echocardiographic assessments, is associated with improved survival and reduced hospitalizations, whereas persistent or progressive pathological remodeling correlates with higher morbidity and mortality.
- Suboptimal adherence to Brazilian Society of Cardiology (SBC) guidelines for heart failure management contributes to preventable adverse outcomes, underscoring the need for local implementation strategies and optimization of ambulatory care.

2 OBJECTIVES

2.1 GENERAL OBJECTIVE

The general objective of this thesis is to describe the clinical, laboratory, and imaging characteristics of ambulatory patients with chronic heart failure (CHF) followed at a Brazilian tertiary academic center (InCor–HCFMUSP), and to evaluate the incidence, risk factors, and prognostic impact of mortality, hospitalization, and electrolyte disturbances in this population.

2.2 SPECIFIC OBJECTIVES

- To assess all-cause mortality rates and identify clinical, laboratory, and therapeutic predictors of mortality among ambulatory CHF patients.
- To describe sociodemographic, clinical, pharmacological, biochemical, and echocardiographic characteristics, as well as healthcare resource utilization, in this ambulatory CHF cohort.
- To evaluate the cumulative incidence and rates of all-cause hospitalizations during follow-up and their relationship with heart failure severity and treatment patterns.
- To estimate the cumulative incidence and rates of mild (K^+ >5.0–5.5 mEq/L), moderate (K^+ >5.5–6.0 mEq/L), and severe (K^+ >6.0 mEq/L) hyperkalemia, and to identify factors associated with the first episode and recurrence of hyperkalemia.
- To compare longitudinal variations in renal function markers (urea, creatinine, eGFR) between baseline and last follow-up, estimating the incidence of renal dysfunction over time.
- To examine adherence to the Brazilian Society of Cardiology Heart Failure Guidelines regarding pharmacological therapy, and to explore the association between adherence level and clinical outcomes.
- To estimate the incidence of reverse and pathological cardiac remodeling using serial echocardiographic parameters, and to investigate their relationship with long-term morbidity and mortality.

3 ARTICLE 1

scientific reports

Check for updates

OPEN

Association of potassium disorders with the mode of death and etiology in patients with chronic heart failure: the INCOR-HF study

Ivna G C V Lima[1], Jairo T Nunes[1], Igor H de Oliveira[2], Silvia M A Ferreira[1], Robinson T Munhoz[1], Paulo R Chizzola[1], Bruno Biselli[1], Brenno R Gomes[1], Lucas P Damiani[3], André S Maria[4], Fernanda Ronco[4] & Edimar Alcides Bocchi[1]✉

Observational studies suggest a U-shaped association between serum potassium (K^+) levels and mortality in patients with chronic heart failure (CHF). However, the mode of death in patients with HF and K^+ disorders remains speculative. To investigate the association between potassium disorders and the mode of death in patients with CHF. A retrospective cohort of 10,378 CHF outpatients was analyzed over an average of 3.28 ± 2.5 years. Kaplan-Meier method, Cox proportional hazards regression models, Poisson regression models adjusting for confounders, and e-value determination ($e' > 1.6$) were used to observe associations between potassium disorders and outcomes. Chagas etiology ($p < 0.01$) and triple HF therapy ($p < 0.01$) were associated with hyperkalemia. Atrial fibrillation was associated with hypokalemia ($p < 0.01$). Chronic kidney disease (CKD) ($p < 0.01$) and diabetes ($p = 0.03$) were associated with both. Hypertension was inversely related to hyperkalemia ($p < 0.01$); age was inversely related to hypokalemia. Associations with mortality were significant for Chagas ($p < 0.01$, e-value 2.16), stroke ($p < 0.01$, e-value 1.85), hypokalemia ($p = 0.02$, e-value 1.94), severe hyperkalemia ($p = 0.08$, e-value 1.93), and CKD ($p < 0.01$, e-value > 1.63). Decompensated HF or cardiogenic shock was the cause of death in 54% of patients with normokalemia, 67.8% with hypokalemia, 44.9% with mild hyperkalemia, 57.8% with moderate hyperkalemia, and 69% with severe hyperkalemia. Most patients with hypokalemia and severe hyperkalemia died from decompensated HF ($p = 0.007$). Data suggest hypokalemia and severe hyperkalemia, along with Chagas and CKD, are associated with death. Unexpectedly, progressive HF was the most frequent mode of death rather than arrhythmias. Further studies are needed to confirm these findings and explore the underlying mechanisms.

Keywords Heart failure, Chagas disease, Hypokalemia, Hyperkalemia, Mortality, Triple therapy

Chronic heart failure (CHF) is a prevalent clinical condition affecting millions globally, characterized by high morbidity and mortality rates. It imposes significant financial burdens on healthcare systems[1]. Potassium disorders, observed in 3 to 18% of patients in randomized clinical trials and up to 25% in observational studies, are linked to adverse outcomes in CHF, with varying impacts depending on the heart failure (HF) phenotype.

Studies have shown that serum potassium levels are independently associated with cardiovascular death, hospitalization due to HF, and heart transplantation in symptomatic CHF patients[2]. In HF with reduced ejection fraction (HFrEF), potassium disorders correlate with HF hospitalizations, with hypokalemia particularly associated with both cardiovascular and non-cardiovascular mortality, especially in patients with chronic kidney disease (CKD). Conversely, in HF with preserved ejection fraction (HFpEF), both hypokalemia and hyperkalemia are linked to increased risks of cardiovascular death, sudden death, and HF death[3]. Hyperkalemia is also a barrier to the prescription of guideline-directed medical therapy (GDMT), leading to increased cardiovascular events.

[1]Heart Failure Clinics, Instituto do Coracao, Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, SP, Brazil. [2]Medical and Hospital Information Division, Instituto do Coracao, Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, SP, Brazil. [3]Academic Research Organization (ARO), Hospital Israelita Albert Einstein, Sao Paulo, Brazil. [4]AstraZeneca, São Paulo, Brazil. ✉email: dcledimar@incor.usp.br

Despite these findings, questions remain about whether potassium disorders specifically increase the risk of arrhythmias or sudden death in CHF patients[4–6]. This study aims to investigate the association between potassium disorders and the mode of death in CHF patients, exploring whether hyperkalemia and hypokalemia contribute differently to this association and examining how hydroelectrolytic disorders influence etiologies prone to sudden death or arrhythmias.

Methods

Study design and patient population

A retrospective cohort study of patients with CHF receiving outpatient care at a tertiary cardiology center between January 2013 and December 2020 was conducted. Data were collected from the electronic medical records system, encompassing patient demographics, laboratory variables (including serum potassium and other electrolytes), imaging tests, comorbidities, medication use, hospitalizations, and emergency room visits. Comorbidities were identified using ICD-10 and procedure codes, with all information verified for accuracy. Clinical data such as functional class (NYHA), weight, blood pressure, heart rate, and CHF etiology were confirmed individually in the medical records. The estimated glomerular filtration rate (eGFR) was calculated using the CKD-EPI formula within six months before 2013 to capture baseline laboratory information.

Inclusion and exclusion criteria

Patients were included if they had at least one serum potassium measurement within six months of the study's start. If this test was conducted before 2013, that date was considered the index consultation. Patients lacking baseline variables in medical records or those participating in other institutional protocols were excluded.

Laboratory tests

Chronic kidney disease (CKD) was classified into five stages according to the eGFR determined by KDIGO guidelines[7], without considering albuminuria due to its low testing frequency in this population. Hyperkalemia was categorized as mild ($5.0 < K^+ < 5.5$ mEq/L), moderate ($5.5 < K^+ < 6.0$ mEq/L), and severe ($K^+ > 6.0$ mEq/L). Hypokalemia was classified as mild ($3.0 < K^+ < 3.5$ mEq/L) and severe ($K^+ < 3.0$ mEq/L).

Outcomes

The primary outcome was mortality from any cause, obtained from the Death Information System (SIM) of the Health Department, adhering to confidentiality and data protection laws, This study was conducted with the approval of the Institutional Ethics Committee of the Heart Institute (InCor) at the Hospital das Clínicas, Faculty of Medicine, University of São Paulo. Mortality records from death certificates were used, with diseases or injuries contributing to death listed hierarchically. The underlying cause of death was identified, and comorbidities contributing to death were reported. Deaths were classified by location (hospital or out-of-hospital) and analyzed according to underlying potassium levels and the last potassium measurement before death. Causes of death were categorized using ICD-10 codes, focusing on cardiogenic shock, decompensated HF, and arrhythmias. All research was conducted in accordance with relevant guidelines and regulations and received ethical approval.

Control measures

Given the retrospective nature of this study, we implemented rigorous statistical controls to mitigate potential biases and confounding factors: To assess the robustness of our findings against unmeasured confounding, we performed an E-value calculation. The E-value quantifies the minimum strength of association that an unmeasured confounder would need to have with both the exposure and the outcome to explain away the observed association fully[8]. We adjusted for known confounders such as age, sex, comorbidities, medication use, and clinical variables.

Statistical analysis

Patient characteristics were summarized using descriptive statistics. Categorical variables were presented as frequencies and percentages, while continuous variables were summarized using means and standard deviations or medians and interquartile ranges, as appropriate. Univariate analyses were conducted to examine the association between baseline variables and the outcomes of interest. Variables considered in the univariate analysis included age, etiology, NYHA functional class, previous stroke, hypertension, diabetes, atrial fibrillation, use of triple therapy, CKD stages, and serum potassium levels. The associations were evaluated using chi-square tests for categorical variables and t-tests or Mann-Whitney U tests for continuous variables. To control for potential confounding factors, multivariate regression models were employed. Cox proportional hazards regression was used to assess the association between serum potassium levels and mortality. Time-to-event data were analyzed using Kaplan-Meier survival curves, and differences between groups were evaluated using the log-rank test. Variables with a p-value < 0.10 in univariate analysis were included in the multivariate Cox regression models. Logistic regression was employed to evaluate the association between potassium disorders (hyperkalemia and hypokalemia) and various covariates. The results were presented as odds ratios (ORs) with 95% confidence intervals (CIs). Given the longitudinal nature of the data, time-dependent Cox regression models were applied to account for changes in exposure status over time. This approach allowed for the assessment of how variations in serum potassium levels over time influenced mortality risk. Missing data were addressed using multiple imputation techniques. Multiple imputed datasets were created, each replacing missing values with plausible values based on observed data. The results from these datasets were combined to produce final estimates that account for the uncertainty associated with the missing data. Mortality rates were calculated using Poisson regression models, presented as deaths per person-year. This approach accommodated varying

follow-up times among patients, providing a standardized mortality rate for comparison. To assess the potential impact of unmeasured confounders, an E-value calculation was performed. The E-value quantifies the minimum strength of association that an unmeasured confounder would need to have with both the exposure (serum potassium levels) and the outcome (mortality) to explain away the observed association[9]. All statistical analyses were conducted using R version 4.2.0[10]. A two-tailed significance level of 5% was applied for all tests, with p-values < 0.05 considered statistically significant.

Handling missing data

Patients without potassium measurements, missing demographic variables, participation in other protocols, or duplicate registrations were excluded. Laboratory variables with less than 20% missing data were adjusted as covariables. Multiple imputations handled missing data by creating multiple imputed datasets, substituting missing values with plausible values based on observed data. These datasets were analyzed separately and combined to account for the uncertainty caused by missing data.

Results

Baseline characteristics

The study sample consisted of patients who were followed in a heart failure clinic, with data collected from their clinical records. Data from 10,423 CHF patients were analyzed, with a mean follow-up of 3.28 ± 2.5 years. Baseline laboratory tests were available for 92% of the patients, with an average serum potassium level of 4.5 ± 0.5 mEq/L. Hypokalemia was present in 1.6% of patients, while 13% had hyperkalemia. Most patients had HF with reduced ejection fraction (HFrEF) and were symptomatic at their first visit. Common comorbidities included coronary artery disease (CAD), dyslipidemia, hypertension, atrial fibrillation (AF), diabetes mellitus (DM), smoking, alcoholism, and acute myocardial infarction (AMI). Chagas disease accounted for 5.3% of cases. As shown in Tables 1 and 80% of the patients had a left ventricular ejection fraction (LVEF) below 50%, indicating a significant proportion of patients with reduced ejection fraction. (Table 1).

For hypokalemia recurrence, the Poisson model indicated associations with valvular etiology (aRR 1.59, 95% CI: 1.07-2.37, $p < 0.01$), AF (aRR 1.71, 95% CI: 1.39-2.10), and moderate to severe CKD (aRR 2.61-6.08, 95% CI: 1.77-6.94, $p < 0.01$). Age was weakly protective against hypokalemia recurrence (Tables 2 and 3).

Chagas etiology was associated with higher mortality risk ($p < 0.01$), hyperkalemia, and recurrence. E-value analysis yielded a value of 2.16, indicating a significant risk associated with this etiology, likely not due to unmeasured factors. Valvular and idiopathic etiologies were weakly associated with hypokalemia. In valvular cardiomyopathy patients, the risk of hypokalemia recurrence was higher ($p < 0.01$).

Outcomes

Out of 10,378 patients, 2,311 died, resulting in a mortality rate of 22.2%. The mortality incidence rate was 7.44 per 100 patient-years (95% CI: 7.16-7.74). Hospitalization incidence was 10.71 per 100 patient-years (95% CI: 10.32-11.11). First hospitalization incidence was 5.08 per 100 patient-years (95% CI: 4.83-5.34). First-time emergency room access incidence was 9.53 per 100 patient-years (95% CI: 9.17-9.9).

Independent mortality risk factors in the Cox-adjusted model included Chagas etiology, age, stroke, hypokalemia, and all levels of hyperkalemia. Adjusting for CKD, hypokalemia and severe hyperkalemia remained significant risk factors, while hypertension was weakly protective (Table 4). The Kaplan-Meier curve (Fig. 1) demonstrated cumulative mortality adjusted for baseline potassium levels and potassium ranges. The multivariate model showed the effect of potassium disorders on mortality: hypokalemia (aHR 1.96, 95% CI: 1.52-2.54, $p < 0.01$), mild hyperkalemia (aHR 1.19, 95% CI: 1.05-1.36, $p = 0.01$), moderate hyperkalemia (aHR 1.33, 95% CI: 1.04-1.70, $p = 0.02$), and severe hyperkalemia (aHR 1.49, 95% CI: 1.0-2.21, $p = 0.05$).

A cubic polynomial model adjusted for age and year of inclusion estimated a reversed J-shaped curve for baseline potassium levels, with the highest death risk at potassium levels ≤ 3 mEq/L and ≥ 6 mEq/L and the lowest risk between 4.3 and 4.6 mEq/L. Sensitivity analysis, including the e-value, is provided in Table 4.

Decompensated HF and cardiogenic shock, followed by ischemic heart disease, were the most frequent causes of death (Table 2). Electrolyte disorders contributed to 1% of deaths, with arrhythmias accounting for 6.2%. Other significant causes included septicemia/infections, stroke, pulmonary embolism, and respiratory failure. Analyzing baseline potassium levels as a predictor of death mode revealed no significant difference between potassium levels and the place of death ($p = 0.475$). However, patients with severe hypokalemia and primary hyperkalemia tended to die from HF-related causes ($p = 0.08$). To enhance the accuracy of a potential association between potassium levels and outcomes, we opted to use the most recent potassium level measured before the event. This approach ensures that the potassium measurements are temporally closer to the outcome, thereby providing a more reliable assessment of the relationship between potassium levels and the event. The last potassium measurement before death indicated that hypokalemia and hyperkalemia were associated with higher risks of death due to HF, cardiogenic shock, or arrhythmia ($p = 0.03$), with these patients often dying from decompensated HF ($p = 0.007$) (Table 5).

Potassium disorders

The incidence rate of hyperkalemia was 22.72 per 100 patient-years (95% CI: 22.01-23.48), while hypokalemia occurred at a rate of 3.56 per 100 patient-years (95% CI: 3.32-3.81). Mild hyperkalemia had an incidence rate of 19.31 (95% CI: 18.66-19.98), moderate hyperkalemia 5.07 (95% CI: 4.78-5.37), and severe hyperkalemia 1.77 (95% CI: 1.6-1.94).

Variables associated with the incidence of hyperkalemia included the Chagas disease, age, DM, use of triple therapy, and all stages of CKD (adjusted hazard ratio [aHR] range from 1.49 to 5.53, $p < 0.01$). Hypertension was inversely related to hyperkalemia (aHR 0.84, 95% CI: 0.76-0.93, $p < 0.01$). Chagas etiology, age, DM, and

Baseline variables	Total (n=10,378)	Normokalemia (n=7866)	Hyperkalemia (n=1193)	Hypokalemia (n=147)	p
Age; mean ± SD	54 ± 13.8 (n=10378)	53.6 ± 13.5 (n=7866)	58.5 ± 12 (n=1193)	55.9 ± 13.4 (n=147)	<0.001
Sex					<0.001
Female	35.2%	36.1%	27.2%	49%	
Male	64%	63.3%	72.3%	50.3%	
Other	0.8%	0.6%	0.5%	0.7%	
etiology					0.370
Ischemic	39.6%	39.5%	42.7%	35.4%	
Chagas	5.3%	5.5%	5.8%	5.6%	
Idiopathic	14.5%	14.6%	13.4%	13.9%	
Valvar heart disease	5.2%	5.3%	4%	6.9%	
Hypertensive	24%	24%	24.3%	25.7%	
Other	11.3%	11.1%	9.7%	12.5%	
Comorbidities					
Atrial fibrillation	29.9%	30.5%	30%	39.5%	0.059
AMI	10.6%	10.4%	12%	6.1%	0.051
CAD	39.2%	39%	42.7%	34.7%	0.024
NYHA					0.140
I	38.1%	37.9%	38.1%	46.9%	
II	39.4%	39.4%	39.2%	27.2%	
III	19.8%	19.8%	19.6%	23.1%	
IV	2.8%	2.8%	3.1%	2.7%	
Deep venous thrombosis	0.9%	0.8%	1.1%	2.7%	0.039
Anemia	4.4%	4.4%	4.2%	6.1%	0.559
Diabetes	29.4%	29.5%	34.3%	25.9%	0.002
Thyroid					0.288
Hyperthyroidism	0.1%	0.2%	0%	0%	
Hypothyroidism	2.8%	2.9%	2.1%	2.7%	
Dyslipidemia	38.1%	40.1%	42.3%	34.7%	0.094
Hypertension	36.2%	36.3%	37.2%	39.5%	0.615
Stroke					0.294
Ischemic	2.4%	2.4%	2.5%	1.4%	
Hemorrhagic	0.6%	0.7%	0.5%	0.7%	
COPD	8.3%	7.9%	8.8%	11.5%	0.031
Smoking	12.5%	12.8%	12.9%	16.3%	0.443
Alcoholism	10.7%	11.1%	11.4%	14.3%	0.450
HIV	0.1%	0.1%	0%	0%	0.551
Valvar disease	2.4%	2.4%	2%	3.4%	0.524
Connective tissue disease	0.3%	0.3%	0.1%	0%	0.374
BMI, kg/m^2	26.9 ± 4.9 (n=8719)	27 ± 4.9 (n=6756)	26.4 ± 4.4 (n=1025)	27.3 ± 5.1 (n=127)	<0.001
DBP, mmHg	75.6 ± 12.9 (n=8388)	75.8 ± 12.9 (n=6541)	74.5 ± 12.9 (n=977)	76.7 ± 15 (n=121)	0.212
SBP, mmHg	119.6 ± 21.1 (n=8388)	119.8 ± 21 (n=6541)	118.5 ± 21.6 (n=977)	120.5 ± 27.3 (n=121)	0.012
HR, bpm	72.8 ± 14.3 (n=8922)	72.9 ± 14.6 (n=6980)	71.6 ± 12.8 (n=1034)	76.3 ± 13.5 (n=128)	0.001
Ejection Fraction (n)	(n=3282)	(n=2800)	(n=64)	(n=418)	
LVEF < 40%	63.8%	63.3%	66.7%	67.2%	0.128
LVEF 40 - 50%	16.0%	15.9%	17.7%	12.5%	
LVEF > 50%	20.2%	20.9%	15.6%	20.3%	
Spironolactone	69.9%	74.4%	66.5%	74.8%	<0.001
Beta-blocker	84.4%	87.4%	91.5%	85.7%	<0.001
RAASi	66.8%	70.1%	69.6%	56.5%	0.002
ARB	26.8%	28.4%	24%	25.9%	0.006
ARNI	1.0%	1.0%	1.1%	0.9%	0.243
Thiazides	36.8%	38.8%	34.5%	65.3%	<0.001
Statin	57.7%	59.7%	66.3%	50.3%	<0.001
Anticoagulants	26.5%	27.9%	26.9%	31.3%	0.490
Continued					

Baseline variables	Total (*n*=10,378)	Normokalemia (*n*=7866)	Hyperkalemia (*n*=1193)	Hypokalemia (*n*=147)	*p*
Ivabradine	2.5%	2.8%	2.3%	1.4%	0.374
Hydralazine	35%	35.6%	44.6%	48.3%	<0.001
Nitrate	32.3%	32.8%	42.3%	42.2%	<0.001
Aspirin	41.7%	42%	51.2%	34%	<0.001
Furosemide	73.5%	76.2%	80.9%	83%	<0.001
Gliclazide	13.1%	13.8%	15.8%	8.8%	0.040
Ferrous sulphate	4%	4%	5.8%	5.4%	0.015
NSAID	0.1%	0.1%	0.1%	0%	0.936
Triple therapy (RAASi + Beta-blocker + Spironolactone)	53%	56.8%	50.5%	48.3%	<0.001
Alternative Therapy (Hydralazine/Nitrate + Beta-blocker + Spironolactone)	9%	9.3%	9.9%	17.5%	0.004
Mean GFR ± SD	64.5 ± 26.3 (*n*=6671)	66.7 ± 25.7 (*n*=5633)	52.3 ± 25.2 (*n*=829)	53.1 ± 26.6 (*n*=108)	<0.001
Normal	16.8%	18.3%	8%	10.2%	<0.001
Mild reduced	39.8%	41.9%	28.8%	26.9%	
Moderate reduced	32.9%	31.2%	42.8%	40.7%	
Severe reduced	7.5%	6.5%	12.9%	15.7%	
Terminal or dialysis	2.9%	2%	6.5%	7.5%	
Mean potassium ± SD	4.5 ± 0.5 (*n*=9206)	4.4 ± 0.4 (*n*=7866)	5.4 ± 0.4	3.2 ± 0.2	<0.001
BNP (*n*=1758)	461±820	472±800	584±1035	581±962	
Haemoglobin	13.6±1.9	13.7±1.8	13.3±2.1	13.1±1.9	
% Lymphocytes	27.6±9.3	28±9.3	25.9±8.8	25.5±9.2	
Troponin-HS (pg/mL)	2303±438	2302±441	2305±433	2324±383	

Table 1. Clinical characteristics in the index consultation. Categorical variables were compared using the chi-square method and the continuous ones using ANOVA ARNI, neprilysin receptor antagonists; ARB, angiotensin II receptor blocker; NSAID, non-steroidal anti-inflammatory drugs; SD, standard deviation; CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; CKD, kidney disease; LVEF, left ventricular ejection fraction; HR, heart rate; Hb, haemoglobin; HIV, human immunodeficiency virus; ACEI, angiotensin conversion enzyme inhibitor; BMI, body mass index; AMI, acute myocardial infarction; NYHA, New York Heart Association functional class; BNP, brain natriuretic peptide; DBP, diastolic blood pressure; SBP, systolic blood pressure; GFR, glomerular filtration rate; DVT, deep vein thrombosis; Troponin-US, highly-sensitive troponin

Variable	Hyperkalemia (*n*=5446) HR (95% CI)	*P*	Hypokalemia (*n*=5446) HR (95% CI)	*P*
Chagas etiology	1.54 (1.26 - 1.89)	<0.01	1.14 (0.77 - 1.68)	0.51
Idiopathic etiology	0.87 (0.75 - 1.01)	0.07	0.81 (0.60 - 1.10)	0.18
Age	1.02 (1.01 - 1.02)	<0.01	0.98 (0.97 - 0.99)	<0.01
Hypertension	0.84 (0.76 - 0.93)	<0.01	1.02 (0.84 - 1.25)	0.81
Diabetes	1.13 (1.02 - 1.25)	0.01	1.23 (1.02 - 1.50)	0.03
Atrial fibrillation	1.03 (0.94 - 1.14)	0.52	1.67 (1.39 - 2.01)	<0.01
Triple therapy[a]	1.34 (1.21 - 1.48)	<0.01	0.86 (0.71 - 1.04)	0.12
Mild reduced CKD	1.49 (1.26 - 1.77)	<0.01	1.76 (1.24 - 2.49)	<0.01
Moderate reduced CKD	2.52 (2.12 - 3.00)	<0.01	3.08 (2.16 - 4.40)	<0.01
Severe reduced CKD	4.07 (3.26 - 5.07)	<0.01	7.76 (5.19 - 11.61)	<0.01
Terminal or dialysis CKD	5.53 (4.08 - 7.51)	<0.01	5.03 (2.81 - 8.99)	<0.01

Table 2. Multivariate Cox model for incidence of potassium disorders. All etiologies were considered, but only those with a significant association were found. CKD, chronic kidney disease; HR, hazard ratio; 95% confidence interval (95% CI). [a]Triple therapy use of RASSi + beta-blocker + spironolactone.

Variable	Hyperkalemia Relative Risk (95% CI)	P	Hypokalemia Relative Risk (95% CI)	P
Chagas etiology	1.35 (1.11 - 1.64)	<0.001	1.3 (0.87 - 1.92)	0.046
Idiopathic etiology	0.84 (0.72 - 0.99)	0.001	0.75 (0.53 - 1.06)	0.011
Valvar heart disease	0.91 (0.72 - 1.16)	0.235	1.59 (1.07 - 2.37)	<0.001
Hypertensive aetiology	0.92 (0.81 - 1.04)	0.041	0.94 (0.71 - 1.23)	0.455
Age (years)	1.01 (1.01 - 1.01)	<0.001	0.98 (0.97 - 0.99)	<0.001
Stroke	1.15 (0.91 - 1.45)	0.062	1.17 (0.74 - 1.84)	0.294
AMI	1.07 (0.92 - 1.25)	0.154	0.82 (0.57 - 1.18)	0.092
Hypertension	0.85 (0.77 - 0.95)	<0.001	1.09 (0.88 - 1.36)	0.196
Diabetes	1.07 (0.97 - 1.19)	0.029	1.14 (0.92 - 1.42)	0.058
Atrial fibrillation	1.06 (0.96 - 1.18)	0.058	1.71 (1.39 - 2.1)	<0.001
Triple therapy	1.21 (1.09 - 1.34)	<0.001	0.91 (0.74 - 1.12)	0.149
Mild reduced CKD	1.64 (1.35 – 2)	<0.001	1.4 (0.96 - 2.05)	0.007
Moderate reduced CKD	2.7 (2.21 - 3.29)	<0.001	2.61 (1.77 - 3.83)	<0.001
Severe reduced CKD	3.94 (3.12 - 4.98)	<0.001	6.08 (3.95 - 9.35)	<0.001
Terminal or dialysis CKD	4.63 (3.47 - 6.18)	<0.001	3.91 (2.2 - 6.94)	<0.001

Table 3. Poisson model for recurrence of potassium disorders. CKD, chronic kidney disease; AMI, acute myocardial infarction; 95% CI, 95% confidence interval.

Models for mortality (*n*=8901)	Multivariate HR (95% CI)	P	E-value
Chagas etiology	1.64 (1.35 - 1.99)	<0.01	2.16
Valvar heart disease	1.25 (1.00 - 1.58)	0.05	1.62
Age, years	1.01 (1.01 - 1.02)	<0.01	1.11
Stroke	1.41 (1.12 - 1.77)	<0.01	1.85
Hypertension	0.88 (0.79 - 0.98)	0.02	1.42
Diabetes	-	-	-
Atrial fibrillation	1.11 (1.00 - 1.23)	0.05	1.36
Use of triple therapy	0.89 (0.80 - 0.99)	0.03	1.39
Hypokalemia	1.47 (1.06 - 2.02)	0.02	1.94
Mild hyperkalaemia	1.05 (0.90 - 1.24)	0.52	1.24
Moderate hyperkalaemia	0.96 (0.70 - 1.31)	0.78	1.21
Severe hyperkalemia	1.47 (0.96 - 2.25)	0.08	1.93
Mild reduced CKD	1.26 (1.04 - 1.53)	0.02	1.63
Moderate reduced CKD	2.01 (1.65 - 2.45)	<0.01	2.61
Severe reduced CKD	3.77 (3.00 - 4.73)	<0.01	4.36
Terminal or dialysis CKD	4.32 (3.26 – 5.72)	<0.01	4.82

Table 4. Multivariable Cox model for mortality and sensitivity analysis. All variables were considered for analysis, but only those with a significant association are presented. CKD, chronic kidney disease; HR, hazard ratio; AMI, acute myocardial infarction; 95% CI%, 95% confidence interval.

use of triple therapy were positively associated with hyperkalemia (aHR 1.54, 1.02, 1.13, and 1.34 respectively, $p < 0.01$ for all) (Table 2).

In the Poisson model for hyperkalemia recurrence, associated variables were Chagas etiology (adjusted relative risk [aRR] 1.35, 95% CI: 1.11-1.64, $p < 0.01$), age (aRR 1.35, 95% CI: 1.11-1.64, $p < 0.01$), DM (aRR 1.35, 95% CI: 1.11-1.64, $p < 0.01$), use of triple therapy, and all stages of CKD (aRR 1.64-4.63, 95% CI: 1.35-6.18, $p < 0.01$) (Table 3).

Factors associated with hypokalemia incidence were DM (aHR 1.23, 95% CI: 1.02-1.50, $p = 0.03$), AF (aHR 1.67, 95% CI: 1.39-2.01, $p < 0.01$), and all stages of CKD (aHR 1.76-7.76, 95% CI: 1.24-11.61, $p < 0.01$). Age was a weak protective factor against hypokalemia incidence (Table 2).

Discussion

This study is the first to describe the association of hypokalemia and severe hyperkalemia with modes of death from heart failure or cardiogenic shock in patients with chronic heart failure. Our findings revealed an E-value greater than 1.6 for mortality associations with Chagas etiology, valvular etiology, stroke, hypokalemia,

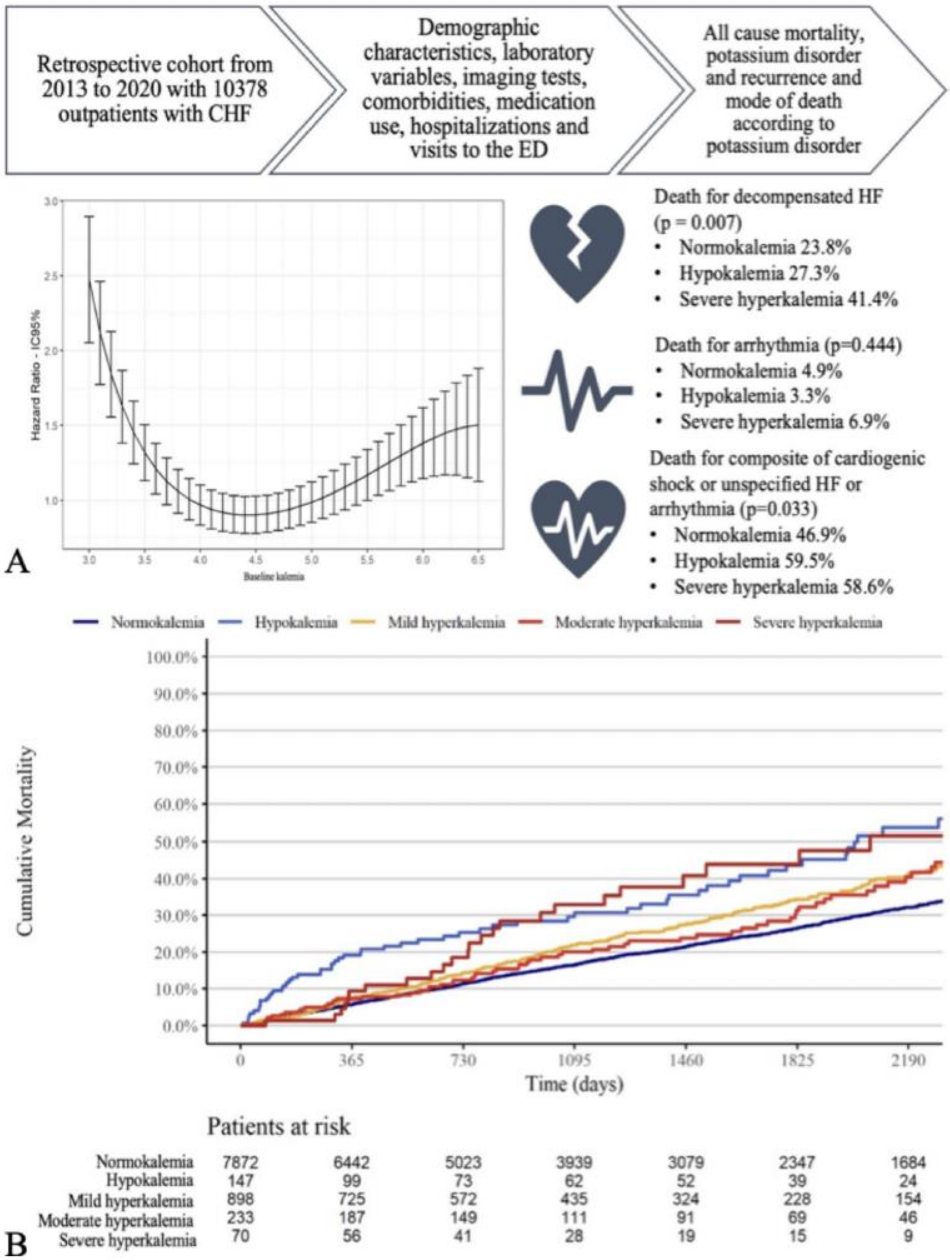


Fig. 1. Structured graphical abstract. **A** Risk curve for mortality according to baseline potassium. **B** Kaplan-Meier of cumulative mortality according to baseline kalaemia.

severe hyperkalemia, and chronic kidney disease (CKD). We observed the influence of specific etiologies on the development and recurrence of potassium disorders: Chagas etiology was linked to a higher incidence and recurrence of hyperkalemia. In contrast, valvar etiologies were more commonly associated with hypokalemia and its recurrence. Additionally, triple therapy, age, diabetes mellitus (DM), Chagas etiology, idiopathic etiology, and CKD were significantly associated with hyperkalemia, whereas hypertension was a protective factor. Conversely, DM, atrial fibrillation (AF), and CKD were associated with hypokalemia, while age served as a protective factor.

Mode of death	Normokalemia (*n*=7866)	Hypokalemia (*n*=147)	Mild hyperkalemia (*n*=898)	Moderate hyperkalemia (*n*=225)	Severe hyperkalemia (*n*=70)	Total (*n*=9206)	*P*
Cardiogenic shock or unspecified HF or arrhythmia	876/1868 (46.9%)	72/121 (59.5%)	105/243 (43.2%)	31/64 (48.4%)	17/29 (58.6%)	1101/2325 (47.4%)	0.033
Cardiogenic shock	564/1868 (30.2%)	49/121 (40.5%)	68/243 (28%)	17/64 (26.6%)	8/29 (27.6%)	706/2325 (30.4%)	0.131
Decompensated HF	444/1868 (23.8%)	33/121 (27.3%)	41/243 (16.9%)	20/64 (31.2%)	12/29 (41.4%)	550/2325 (23.7%)	0.007
Arrhythmias	91/1868 (4.9%)	4/121 (3.3%)	11/243 (4.5%)	6/64 (9.4%)	2/29 (6.9%)	114/2325 (4.9%)	0.444
Hospital death and other health establishment	1477/1687 (87.5%)	48/50 (96%)	209/238 (87.8%)	57/61 (93.4%)	22/23 (95.6%)	2033/2311 (88%)	0.475
Death at home or outside of hospital	189/1687 (11.3%)	2/50 (4%)	29/238 (12.2%)	4/61 (6.6%)	1/23 (4.4%)	250/2311 (10.8%)	

Table 5. Mode of death divided by the last kalaemia before death and local of death.

Previous studies have identified risk factors for hyperkalemia and cardiovascular mortality, including CKD, diabetes, and medication use[11,12]High-dose loop diuretics, CKD, and neurohormonal activation have been associated with hypokalemia and death. Our study found a higher incidence of mild to severe hyperkalemia (22%) and hypokalemia (3.5%) compared to previous reports[13,14], likely due to a higher percentage of patients in NYHA III/IV functional class, renal dysfunction, diuretic use, and RAASi use. Furthermore, we identified Chagas and valvular etiologies as novel risk factors for potassium disorders in this population.

Contrary to the traditional view that arrhythmias predominantly cause death in patients with potassium disorders, our findings indicate that HF or cardiogenic shock was the most frequent mode of death in patients with hypokalemia and severe hyperkalemia. Both hyperkalemia and hypokalemia were associated with a higher risk of hospitalization for HF and death from HF compared to normokalemia.

The mechanisms underlying the association between potassium disorders and death from HF and cardiogenic shock remain speculative. Potassium disorders may serve as markers of greater disease severity in CHF[15]. Hypokalemia might indicate disease progression, associated with more frequent diuretic use and insufficient neurohormonal inhibition, leading to increased activation of angiotensin II, aldosterone, and norepinephrine[16]. Severe hemodynamic disturbances may cause a catecholamine-induced drop in serum potassium[15]. Hyperkalemia, resulting from decreased potassium elimination, could be a marker of cardiorenal syndrome and other comorbidities like Chagas disease and older age, which were identified as risk factors for mortality[17–19]. Alternatively, potassium disorders could contribute to increased mortality through their effects on myocardial function, the cardiovascular system, neurohormonal activation, coagulation, and endothelial function[20–23]. Potassium can stimulate aldosterone production independently of the renin-angiotensin-aldosterone system and reduce angiotensin II-induced vasoconstriction, especially when there is low nitric oxide availability[24]. Animal and human studies suggest that low serum potassium may increase thrombosis rates, endothelial dysfunction, and platelet aggregation[25]. Additionally, serum potassium levels may influence cardiac systolic function by affecting the Na+/K+ ATPase pump, essential for maintaining cellular ion balance and myocardial contractility[26].

In agreement with these mechanisms, triple therapy medications that increase serum potassium levels and improve survival in HFrEF were also associated with hyperkalemia in our study. The suspension or dose reduction of triple therapy might explain increased mortality, although sub-analyses of the PARAGON-HF and PARADIGM-HF studies suggest that hyperkalemia-related mortality occurred despite constant sacubitril-valsartan doses[5,6].

The association between Chagas etiology and potassium disorders is significant given its prevalence in endemic regions and higher mortality compared to other CHF etiologies. Potassium disorders have not been previously identified as contributors to high mortality in Chagas disease, necessitating further research to elucidate the pathophysiological mechanisms involved[27,28].

Limitations

This retrospective observational study has limitations such as selection bias, confounding, and misclassification. However, it addresses important clinical practice issues not evaluable in clinical trials. We included various patient characteristics to reduce confounding bias and performed sensitivity analyses and e-value calculations to mitigate limitations. The database reflects real-world practice standards, limiting misclassification by reducing human interference in data capture and implementing quality control measures[29].

While the e-value has limitations, it was used to assess causality in observational studies. Retrospective data may be incomplete and non-standardized, but our missing data rate was below 20%, with all patients having at least one potassium measurement. Clinical decisions determine the variables obtained, reflecting real-world practice but potentially limiting the accuracy of certain questions. The use of death certificates to determine the mode and cause of death may introduce inaccuracies, but they provide the most accurate information closest to the time of death. The sample size was large, adjustments were made for measured confounders, and sensitivity analyses were performed for unmeasured confounders, enhancing internal validity and generalizability of results.

Conclusion

Our study underscores the importance of monitoring potassium levels in CHF patients, as both severe hypokalemia and hyperkalemia were associated with increased mortality risk. The most common causes of death related to potassium disorders were the progression of HF and cardiogenic shock. Potassium disorders may serve as therapeutic targets for managing CHF, especially in patients with Chagas, valvar disease, elderly, diabetic, chronic kidney disease, and those on triple therapy. Close monitoring and management of potassium disorders could help identify patients with a high risk of death from decompensated HF or cardiogenic shock. Further research is needed to confirm these findings and explore the underlying mechanisms.

Data availability

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

Received: 29 May 2024; Accepted: 30 September 2024
Published online: 04 December 2024

References

1. Chioncel, O. et al. Epidemiology and one-year outcomes in patients with chronic heart failure and preserved, mid-range and reduced ejection fraction: an analysis of the ESC Heart failure Long-Term Registry. *Eur. J. Heart Fail.* **19** (12), 1574–1585. https://doi.org/10.1002/ejhf.813 (2017).
2. Urbich, M. et al. A systematic review of medical costs associated with heart failure in the USA (2014–2020). *PharmacoEconomics*. **38**, 1219–1236. https://doi.org/10.1007/s40273-020-00952-0 (2020).
3. Ferreira, J. P. et al. Abnormalities of potassium in heart failure: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **75** (22), 2836–2850. https://doi.org/10.1016/j.jacc.2020.04.021 (2020).
4. Toledo, C. C. et al. Serum potassium levels provide prognostic information in symptomatic heart failure beyond traditional clinical variables. *ESC Heart Fail.* **8** (3), 2133–2143. https://doi.org/10.1002/ehf2.13295 (2021). Epub 2021 Mar 18. PMID: 33734611; PMCID: PMC8120348.
5. Ferreira, J. P. et al. Serum potassium in the PARADIGM-HF trial. *Eur. J. Heart Fail.* **22** (11), 2056–2064. https://doi.org/10.1002/ejhf.1987 (2020).
6. Ferreira, J. P. et al. Serum potassium and outcomes in heart failure with preserved ejection fraction: A post-hoc analysis of the PARAGON-HF trial. *Eur. J. Heart Fail.* **23** (5), 776–784. https://doi.org/10.1002/ejhf.2134 (2021).
7. Ketteler, M. et al. Executive summary of the 2017 KDIGO Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD) guideline update: what's changed and why it matters. Kidney Int. ;92(1):26–36. (2017). https://doi.org/10.1016/j.kint.2017.04.006. Erratum in: Kidney Int. 2017;92(6):1558.
8. VanderWeele, T. J. & Ding, P. Analysis in observational research: Introducing the E-value. *Ann. Intern. Med.* **167**, 268–274 (2017).
9. Haneuse, S., VanderWeele, T. J. & Arterburn, D. Using the E-value to assess the potential effect of unmeasured confounding in observational studies. *JAMA*. **321**, 602–603 (2019).
10. Mathur, M. B., Ding, P., Riddell, C. A., & Vander Weele, T. J. Website and R package for computing E-values. *Epidemiology*. **29**, e45–e47 (2018).
11. Henrysson, J., Thunström, E., Chen, X., Fu, M. & Basic, C. Hyperkalaemia as a cause of undertreatment with mineralocorticoid receptor antagonists in heart failure. *ESC Heart Fail.*https://doi.org/10.1002/ehf2.14137 (2022).
12. James, G. et al. Serum potassium variability as a predictor of clinical outcomes in patients with cardiorenal disease or diabetes: A retrospective UK database study. *Clin. Kidney J.* **15** (4), 758–770 (2022).
13. Aldahl, M. et al. Associations of serum potassium levels with mortality in chronic heart failure patients. *Eur. Heart J.* **38** (38), 2890–2896 (2017).
14. Cooper, L. B. et al. Association between potassium level and outcomes in heart failure with reduced ejection fraction: A cohort study from the Swedish Heart failure Registry. *Eur. J. Heart Fail.* **22** (8), 1390–1398. https://doi.org/10.1002/ejhf.1757 (2020).
15. Núñez, J. et al. Long-term potassium monitoring and dynamics in heart failure and risk of mortality. *Circulation*. **137** (13), 1320–1330. https://doi.org/10.1161/CIRCULATIONAHA.117.030576 (2018).
16. Linde, C. et al. Serum potassium and clinical outcomes in heart failure patients: Results of risk calculations in 21 334 patients in the UK. *ESC Heart Fail.* **6** (2), 280–290. https://doi.org/10.1002/ehf2.12402 (2019).
17. Moriyama, H., Kohno, T. & Kohsaka, S. Letter regarding the article 'Effects of hyperkalaemia and non-adherence to renin-angiotensin-aldosterone system inhibitor therapy in patients with heart failure in Italy: A propensity-matched study'. *Eur. J. Heart Fail.* **23** (3), 495–496. https://doi.org/10.1002/ejhf.2081 (2021).
18. Srivastava, T. N. & Young, D. B. Impairment of cardiac function by moderate potassium depletion. *J. Card Fail.* **1**, 195–200 (1995).
19. Desai, A. S. et al. *Eur. Heart J.* ;**36**(30):1990–1997. (2015).
20. Sudhir, K., Kurtz, T. W., Yock, P. G., Connolly, A. J. & Morris, R. C. Potassium preserves endothelial function and enhances aortic compliance in Dahl rats. *Hypertension*. **22**, 3 (1993).
21. Parksook, W. W. & Williams, G. H. Aldosterone and cardiovascular diseases. *Cardiovasc. Res. 2022 Apr* **7**:cvac027. https://doi.org/10.1093/cvr/cvac027. Epub ahead of print.
22. Taddei, S. et al. Effect of potassium on vasodilation to acetylcholine in essential hypertension. *Hypertension*. **23**, 485–490 (1994).
23. Young, D. B., Lin, H. & McCabe, R. D. Potassium's cardiovascular protective mechanisms. *Am. J. Physiol.* **268**, R825–R837 (1995).
24. Alper, A. B. et al. A propensity-matched study of low serum potassium and mortality in older adults with chronic heart failure. *Int. J. Cardiol.* **137** (1), 1–8. https://doi.org/10.1016/j.ijcard.2008.05.047 (2009).
25. Shapiro, J. I., Banerjee, A., Reiss, O. K. & Elkins, N. Acute and chronic hypokalemia sensitize the isolated heart to hypoxic injury. *Am. J. Physiol.* **274** (5), H1598–H1604. https://doi.org/10.1152/ajpheart.1998.274.5.H1598 (1998).
26. Barri, Y. M. & Wingo, C. S. The effects of potassium depletion and supplementation on blood pressure: A clinical review. *Am. J. Med. Sci.* **314**, 37–40 (1997).
27. Issa, V. S. et al. The course of patients with Chagas heart disease during episodes of decompensated heart failure. *ESC Heart Fail.* **8** (2), 1460–1471. https://doi.org/10.1002/ehf2.13232 (2021).
28. Mocelin, A. O. et al. The influence of aetiology on inflammatory and neurohumoral activation in patients with severe heart failure: A prospective study comparing Chagas' heart disease and idiopathic dilated cardiomyopathy. *Eur. J. Heart Fail.* **7** (5), 869–873. https://doi.org/10.1016/j.ejheart.2004.10.014 (2005).
29. Boyko, E. J. Observational research — opportunities and limitations. *J. Diabetes Complicat.* **27** (6), 642–648 (2013).

Acknowledgements

The authors would like to express their sincere gratitude to all the participants who generously contributed

their time and effort to this study. We would also like to acknowledge the support and resources provided by AstraZeneca. We extend our appreciation to the Instituto do Coração staff and the InCor -HF research team for their invaluable assistance in data collection and analysis. Special thanks go to Professor Edimar Bocchi for their valuable insights and expertise. Lastly, we thank our families and loved ones for their unwavering support throughout this research endeavor.

Author contributions

I.G.C.V.L. and J.T.N. conceptualized, designed the study, wrote the initial manuscript, and collected the data. I.H.O collected the data. L.P.D performed the statistical analysis, S.M.A.F., and R.T.M. P.R.C., B.B., B.R.G., F.R., A.S., F.R, and E.A.B. critically reviewed the final manuscript and provided significant intellectual input. All authors reviewed and approved the final manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-74928-x.

Correspondence and requests for materials should be addressed to E.A.B.

4 ARTICLE 2

Machine Learning Phenotypes From a 10,376-Patient HFrEF Cohort

Contextual Overview

HFrEF is a heterogeneous clinical syndrome. Traditional risk stratification relies on isolated variables such as LVEF, BNP, or etiology, which fail to capture multidimensional complexity. This study applies modern **machine-learning phenomapping** to one of the largest real-world HFrEF cohorts worldwide, providing deeper insights into latent clinical structures.

Study Population and Feature Processing

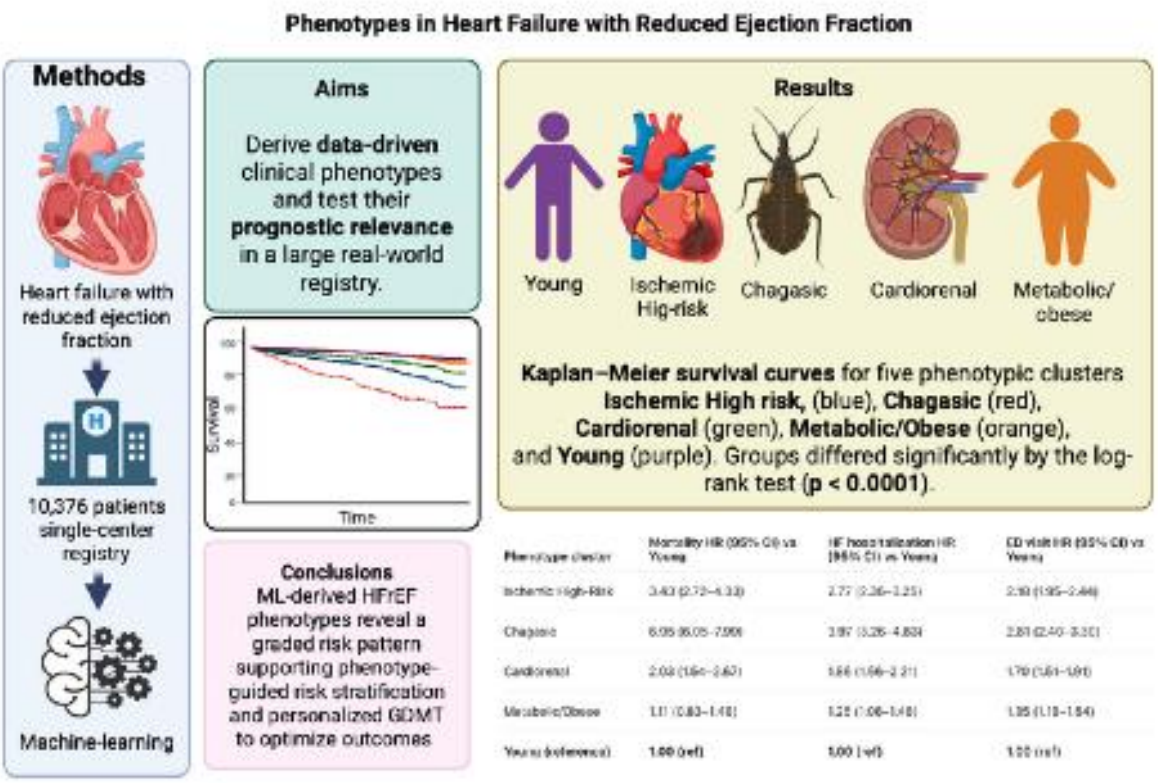
We analyzed **10,376 ambulatory HFrEF patients** from a specialized tertiary HF clinic over a seven-year period. After preprocessing through MICE and dimensionality reduction with FAMD, unsupervised clustering via k-means identified **five reproducible phenotypes**, each representing a distinct disease trajectory within chronic HFrEF

Spline-based visualization of principal components and bootstrap stability analyses (adjusted Rand index) confirmed the structural robustness of the phenotypic solution (**Figure 1**).

Phenotype Descriptions

Each phenotype synthesized demographic, clinical, comorbidity, hemodynamic, renal, and therapeutic dimensions

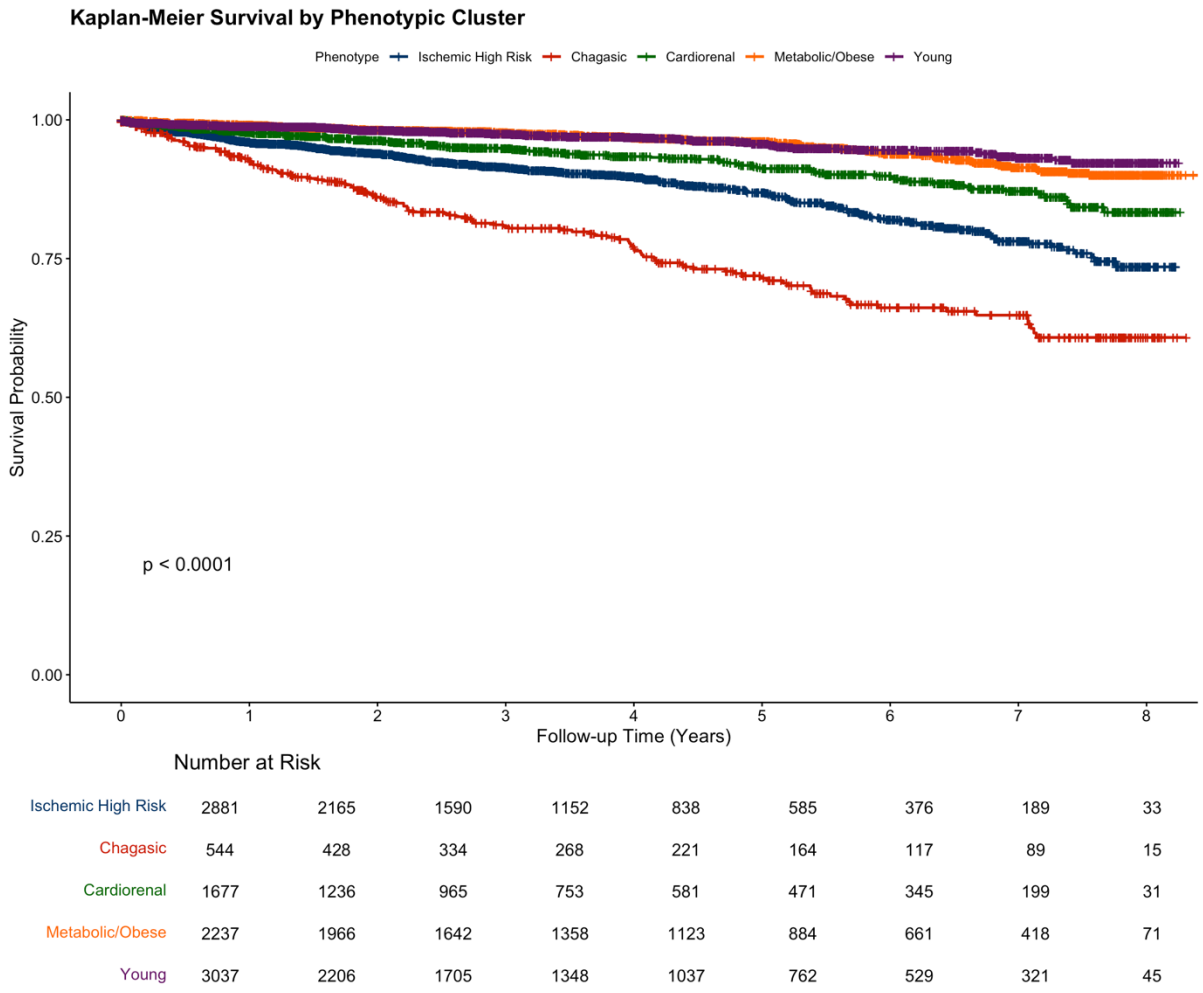
- **Young phenotype:** Characterized by low comorbidity burden, preserved renal function, and predominantly non-ischemic etiologies. Served as the **reference** for prognostic comparisons.
- **Ischemic High-Risk:** Patients were older, overwhelmingly male, with advanced CAD, severe LV dysfunction, high BNP, and extensive GDMT. Despite treatment, their structural disease burden translated into high mortality (9.4%).
- **Chagasic phenotype:** Represented a uniquely severe inflammatory-dilated conduction-system phenotype. Despite a younger age, this group exhibited by far the worst outcomes: mortality 24%; HF hospitalization 31%. This highlights the disproportionate impact of Chagas cardiomyopathy on Brazilian HF populations.
- **Cardiorenal phenotype**: The defining feature was renal dysfunction. Despite mid-range LVEF, this group showed substantial AF prevalence and worse outcomes than expected from LV function alone.
- **Metabolic/Obese:** Patients had high BMI and diabetes prevalence, with preserved renal function and intermediate BNP levels. Mortality was low (4.4%), but ED visits were high (26%), suggesting symptom instability without end-organ compromise.



Mortality Outcomes

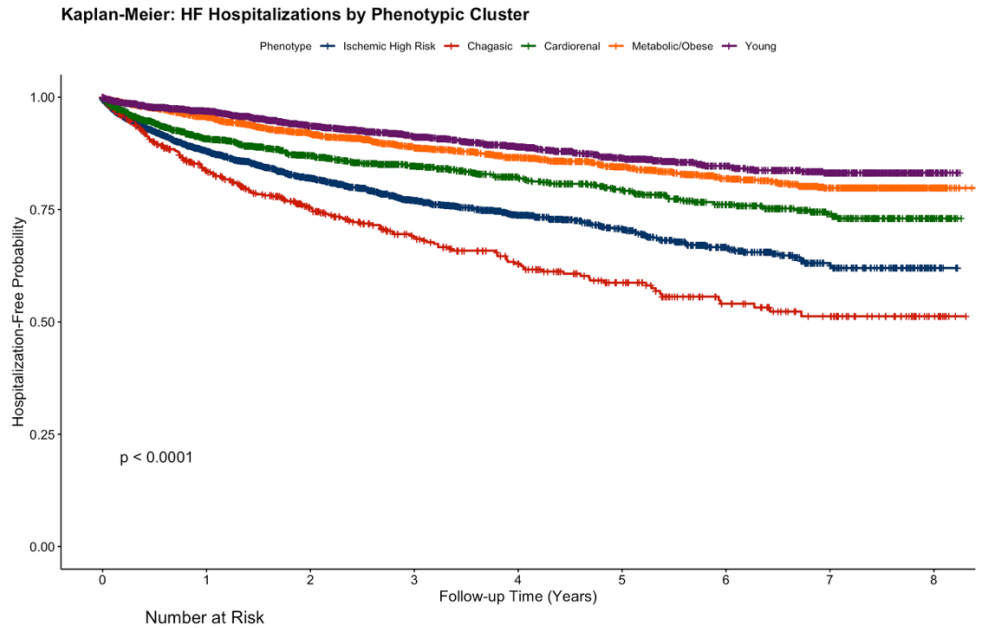
Kaplan–Meier curves **(**bellow**)** demonstrated clear stratification. Versus the Young phenotype:

- Chagasic: HR 6.95
- Ischemic High-Risk: HR 3.43
- Cardiorenal: HR 2.03
- Metabolic/Obese: HR 1.11



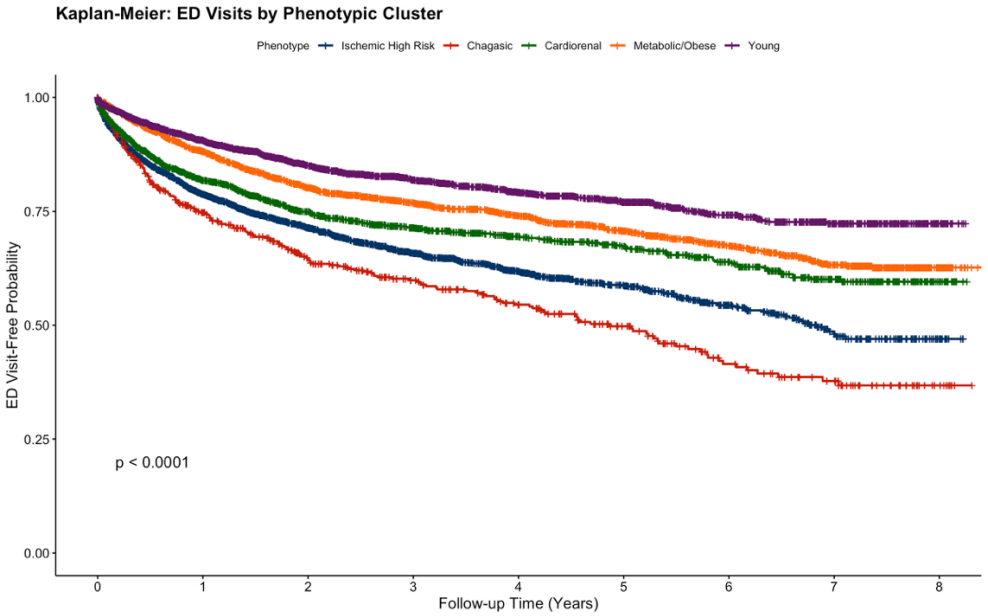
HF Hospitalization and ED Visits

HF hospitalizations followed an identical stepwise gradient. ED visits were disproportionately higher in: Ischemic High-Risk and Chagasic



Number at Risk

Ischemic High Risk	2881	1934	1344	923	640	432	250	116	14
Chagasic	544	369	265	192	149	105	68	47	10
Cardiorenal	1677	1142	861	656	493	389	280	156	27
Metabolic/Obese	2237	1892	1535	1218	991	762	557	349	54
Young	3037	2146	1606	1233	926	667	456	268	33

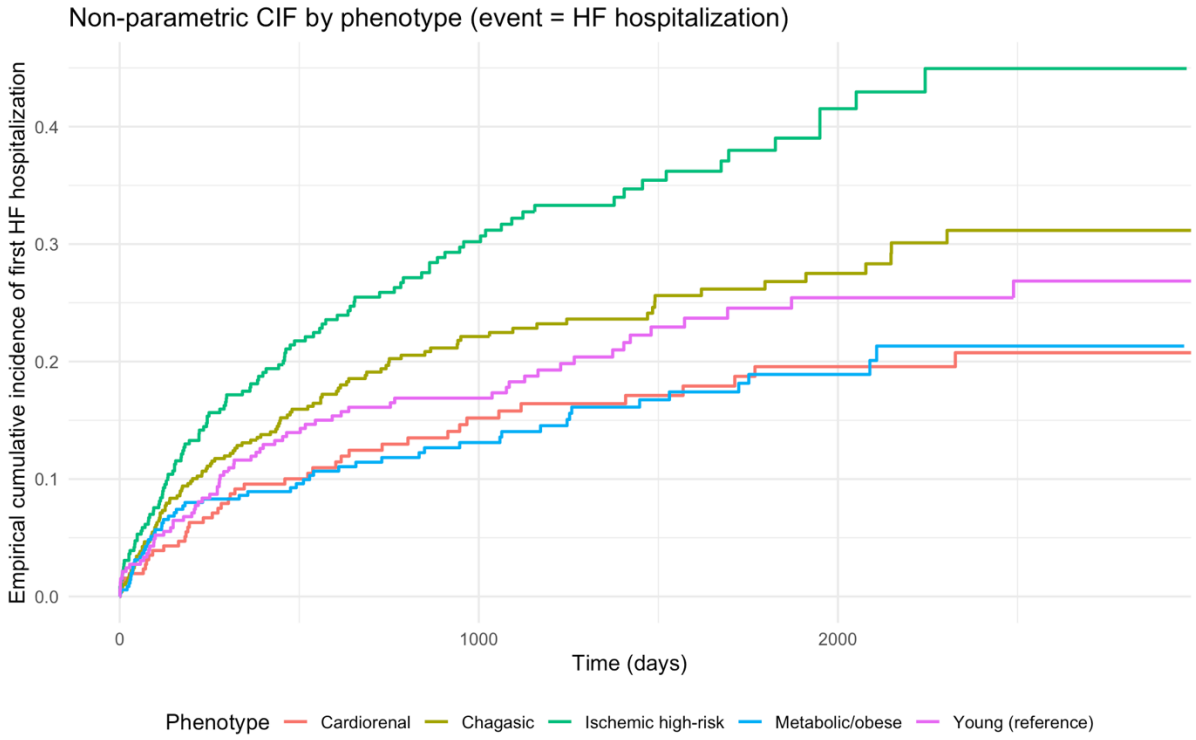


Number at Risk

Ischemic High Risk	2881	1735	1174	799	556	370	212	93	12
Chagasic	544	336	239	181	141	97	61	42	8
Cardiorenal	1677	1034	747	570	426	334	237	126	21
Metabolic/Obese	2237	1746	1349	1070	868	649	468	278	42
Young	3037	2007	1456	1094	810	571	371	211	29

Competing-Risk Analyses

Fine–Gray models revealed attenuated differences, suggesting that **some hospitalization risk is masked by competing early mortality**, especially in Chagas and Ischemic groups.



Interpretation

These findings demonstrate that **machine-learning phenotyping identifies clinically meaningful subtypes** with strong prognostic implications, especially for:

1. risk stratification
2. personalized follow-up intensity
3. targeted GDMT approaches
4. resource allocation in high-risk subgroups

BMJ Open

Machine learning phenotypes in heart failure with reduced ejection fraction from a 10,376-patient real-world cohort

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2025-114674
Article Type:	Original research
Date Submitted by the Author:	29-Nov-2025
Complete List of Authors:	Lima, Ivna; USP Incor, Heart Failure Nunes, Jairo; Peter Munk Cardiac Centre, Heart Failure Rizerio, Brenno; HCFMUSP, InCor Chizzola, Paulo Roberto; Universidade de São Paulo Ayub-Ferreira, Silvia; HCFMUSP, InCor Munhoz, Robinson ; HCFMUSP, InCor Biselli, Bruno; HCFMUSP Ronco, Fernanda; AstraZeneca Pharmaceuticals LP Bocchi, Edimar; HCFMUSP, InCor
Keywords:	Heart failure < CARDIOLOGY, Machine Learning, Prognosis

Abstract

Background: Heart failure with reduced ejection fraction (HFrEF) is heterogeneous, complicating therapy and prognosis. We aimed to derive clinical phenotypes from real-world data and evaluate their prognostic relevance.

Methods: We analysed 10,376 ambulatory patients with HFrEF (2013–2020) from a single tertiary centre registry. Mixed continuous and categorical variables were handled using factorial analysis of mixed data followed by k-means clustering. The primary outcome was all-cause mortality; secondary outcomes were time to first heart failure hospitalisation and time to first emergency department (ED) visit for acute heart failure. Time-to-event associations between phenotypes and outcomes were assessed using Cox proportional hazards models. Fine–Gray competing risk models for first heart failure hospitalisation were performed as sensitivity analyses.

Results: Five phenotypes emerged—Young, Ischaemic High-Risk, Chagasic, Cardiorenal, and Metabolic/Obese—with distinct risk profiles. Compared with the Young group, mortality was highest in the Chagasic (HR, 6.95; 95% CI, 6.05–7.99), Ischaemic High-Risk (HR, 3.43; 95% CI, 2.72–4.33), and Cardiorenal (HR, 2.03; 95% CI, 1.54–2.67) clusters. The Metabolic/Obese group showed intermediate outcomes. Fine–Gray models confirmed attenuated differences for first hospitalisation after accounting for death as a competing event.

Conclusions: Data-driven phenotyping showed distinct and graded differences in mortality, hospitalisation and emergency visit risk across phenotypes. Phenotype-aware strategies may refine prognostic assessment and inform tailored therapeutic interventions, especially in high-risk Chagasic and cardiorenal subgroups.

What is already known on this topic

Heart failure with reduced ejection fraction (HFrEF) is a heterogeneous syndrome, and current risk scores only partially capture patient complexity. Data-driven phenotyping has been proposed as a strategy to refine prognosis and individualise follow-up, but evidence from large, real-world cohorts is limited.

What this study adds

Using unsupervised machine learning in a cohort of 10,376 patients with chronic HFrEF from a specialised centre, we identified five clinically recognisable phenotypes—Young, Chagasic, Ischaemic high-risk, Cardiorenal and Metabolic/obese—with distinct profiles and consistent gradients in mortality, heart failure hospitalisation and emergency department visits.

How this study might affect research, practice or policy

These phenotypes may help clinicians recognise high-risk trajectories using routinely available variables, support phenotype-aware follow-up strategies, and guide the design of future trials and implementation studies targeting specific heart failure subgroups.

Introduction

Heart failure with reduced ejection fraction (HFrEF) remains a major global challenge, associated with high morbidity and recurrent hospitalisations despite advances in therapy.[1–3] Challenges in HF management are driven by disease progression but also by suboptimal implementation of guideline-directed medical therapy (GDMT). Real-world adherence to GDMT recommendations is poor. The GDMT low prescription rate is assigned to patient heterogeneity, patient multiple comorbidities, medical therapeutic inertia, concerns regarding drug side effects, and clinical patient characteristics like hypotension, renal dysfunction, and hyperkalaemia, which complicate treatment intensification. [4,5]

Prior classifications of HFrEF have primarily relied on expert consensus or trial-based subgrouping,[6,7] often excluding high-risk or underrepresented populations such as those with advanced renal dysfunction or Chagas cardiomyopathy. Data-driven clustering offers the opportunity to identify clinically meaningful phenotypes from real-world data, thereby enhancing precision medicine.

This study aims to identify data-driven phenotypes of HFrEF using machine learning-based clustering, assess their prognostic implications, and evaluate their associations with all-cause mortality, hospitalisations, and emergency department visits. By leveraging a phenotyping approach, we aim to refine the understanding of HFrEF heterogeneity and explore its potential to guide personalised therapeutic strategies.[8]

Methods

Study Design and Population

We conducted a retrospective cohort study using a single-centre heart failure registry that includes ambulatory patients managed at a tertiary heart failure clinic. Adults (≥18 years) with left ventricular ejection fraction (LVEF) <50% and at least one outpatient visit between 2013 and 2020 were eligible. The study was approved by the institutional ethics committee, with waiver of informed consent for de-identified

data. Patients or the public were not involved in the design, conduct, reporting or dissemination plans of this research. The study used de-identified data from routine clinical care in a specialised heart failure centre.

Data Collection and Preprocessing

Demographic, clinical, laboratory, imaging, and medication data were extracted from the electronic medical record and supplemented by ICD-10 diagnostic and procedure codes. When possible, baseline laboratory values were obtained within 6 months of the index outpatient encounter.

Patients with more than 20% missing data were excluded from the derivation cohort. Missing data were addressed using multiple imputation via chained equations (MICE). MICE was chosen over simpler techniques, such as mean/mode imputation or k-nearest neighbours, because it preserves the underlying variance structure and prevents artificial inflation of associations—an important consideration when downstream analyses include correlation-sensitive methods, like principal component–based reduction. Both continuous and categorical variables were included in the imputation model. Variables used in clustering were imputed prior to analysis; outcome variables (mortality, hospitalisation, and emergency department visits) were excluded from both imputation and clustering to prevent information leakage. To reduce dimensionality while incorporating both numerical and categorical inputs, we performed factorial analysis of mixed data (FAMD). This technique is designed explicitly for datasets containing both continuous and categorical variables, enabling the joint modelling of diverse clinical inputs, including laboratory values, comorbidities, demographics, and therapy use. Unlike standard principal component analysis (PCA), which assumes all inputs are continuous, FAMD maintains the integrity of categorical structure while preserving the distance geometry needed for clustering. Principal components explaining more than 80% of the total variance were retained and used as input features for unsupervised learning.[10] De-identified data underlying the findings are available from the corresponding author upon reasonable

request, subject to institutional approvals. Custom R code used for preprocessing, FAMD, clustering, and competing-risk analyses is available from the corresponding author upon reasonable request.

Clustering Analysis

Unsupervised clustering was performed using k-means with Euclidean distance. The optimal number of clusters (k = 5) was determined using the Elbow and Silhouette methods and confirmed by expert review. Bootstrap resampling (n = 1000) and the adjusted Rand index were used to assess structural stability. Importantly, no outcome variables were used in the construction or refinement of the clusters. The investigation conforms with the principles outlined in the Declaration of Helsinki.

Outcomes and Validation

The primary outcome was all-cause mortality. Secondary outcomes were time to first heart failure hospitalisation and time to first emergency department (ED) visit for acute heart failure. Vital status and hospitalisation data were obtained from institutional records and national databases. Patients were followed from the baseline assessment until the first occurrence of each outcome, the last contact, or the end of follow-up, whichever came first. Cause-specific Cox models were used in the main analyses. In sensitivity analyses, Fine–Gray subdistribution models accounted for competing risk of death for non-fatal outcomes; these results are reported in the online supplementary material. Internal validation included bootstrap resampling and cluster stability via adjusted Rand index; no external validation cohort was used. The proportional hazards assumption was verified using Schoenfeld residuals. The best prognostic cluster was used as the reference group for all hazard ratio (HR) comparisons.

Statistical analysis

Continuous variables are presented as mean ± SD or median [IQR], and categorical variables as counts (percentages). Between-group comparisons were performed using ANOVA or Kruskal–Wallis tests for continuous variables and χ^2 tests for categorical variables. Missing data were handled using multiple

imputation by chained equations (MICE), with predictive mean matching for continuous and polytomous regression, and for categorical variables, generating 10 imputed datasets. Clustering was conducted on the first imputed dataset for interpretability and replicated across imputations to verify stability.

To identify latent clinical phenotypes, a factorial analysis of mixed data (FAMD) was applied to continuous and categorical variables, followed by k-means clustering. The optimal number of clusters was determined based on the elbow method and silhouette indices, clinical interpretability, and bootstrap stability (adjusted Rand index). Sensitivity analyses using partitioning around medoids (PAM) and Gaussian mixture models (GMM) confirmed the robustness of the clustering solution. Associations between phenotypes and outcomes were evaluated using time-to-event models. Associations between phenotypes and outcomes were evaluated using time-to-event models. Cox proportional hazards models were used for all-cause mortality, first heart failure hospitalisation and first emergency department (ED) visit. In sensitivity analyses for first heart failure hospitalisation, Fine–Gray subdistribution hazard models with death as a competing event were fitted, and cumulative incidence functions were derived by phenotype;

Multivariable models were adjusted for clinically relevant covariates, including age, sex, left ventricular ejection fraction (LVEF), estimated glomerular filtration rate (eGFR), systolic blood pressure (SBP), BNP levels, and baseline use of triple guideline-directed medical therapy (ARNI/ACEI or ARB, MRA, and β-blocker). The Young phenotype was used as the reference category for all models. Proportional hazards assumptions were verified using Schoenfeld residuals. Model robustness was further evaluated through internal bootstrap validation (n = 1000 resamples) and by repeating the clustering analysis after removing redundant variables (variance inflation factor > 5). All analyses were conducted in R version 4.3.3 (R Foundation for Statistical Computing, Vienna, Austria), using the packages *FactoMineR*, *cmprsk*, *riskRegression*, and *ConsensusClusterPlus*. Statistical significance was set at a two-sided $P < 0.05$.[11]

Reporting guidelines

This observational cohort study was reported in accordance with the STROBE recommendations. Key elements related to model development and prognostic profiling were aligned with the TRIPOD guidance where applicable. Completed checklists are provided in the online supplemental material.

Results

Phenotypic Profiles

Unsupervised clustering identified five clinically distinct phenotypes that reflected recognisable trajectories of heart failure (Table 1 and supplementary material). Each phenotype exhibited unique demographic, hemodynamic, and comorbidity patterns with clear clinical plausibility.

The Ischaemic High-Risk phenotype encompassed older patients (median 58 years), predominantly male, with extensive coronary artery disease, low LVEF (27%), and moderate renal dysfunction. Despite the highest use of triple therapy (72%), this group displayed elevated BNP levels and the second-highest mortality (9.4%), representing a classic ischaemic and advanced HFrEF phenotype with suboptimal hemodynamic reserve.

The Chagasic cluster was defined by the near-exclusive presence of Chagas cardiomyopathy (99%), younger age (median 56 years), and markedly adverse outcomes, with a 24% mortality rate and 31% HF hospitalisation. These patients had limited GDMT exposure and frequently presented with low systolic blood pressure and severe right-sided congestion. This subgroup represents a high-risk, inflammatory dilated phenotype that remains underrepresented in clinical trials.

The Cardiorenal cluster included the oldest patients (median age 63 years), with the lowest eGFR (45 mL/min/1.73 m^2), a high prevalence of atrial fibrillation (48%), and a mid-range LVEF. Despite low GDMT adherence, outcomes were poor (mortality 6.3%, HF hospitalisation 15%), reflecting the burden of renal dysfunction as a key determinant of prognosis.

The Metabolic/obese phenotype was characterised by the highest prevalence of obesity (BMI 30.1 kg/m²) and diabetes mellitus (75%), with preserved renal function and modest elevation of BNP. Although mortality (4.4%) and HF hospitalisation (13%) were lower than in other groups, these patients exhibited frequent ED visits (26%), suggesting a metabolic phenotype prone to recurrent decompensation despite preserved perfusion.

Finally, the Young phenotype, comprising the youngest patients (median 47 years) with fewer comorbidities and better kidney function (eGFR 85 mL/min), showed the most favourable prognosis, with the lowest mortality (3.0%) and HF hospitalisation (8%). This cluster likely represents early-stage or genetically driven HFrEF with preserved therapeutic response.

Outcomes

During a median follow-up of 3.2 years, distinct prognostic gradients were observed across the five machine-learning phenotypes. Overall mortality was 6.7%, with marked differences between groups. Compared with the Young phenotype, which served as the reference, the risk of death was significantly higher among the Chagasic cluster (HR 6.95, 95% CI 6.05–7.99; P<0.001), followed by the Ischaemic High-Risk cluster (HR 3.43, 95% CI 2.72–4.33; P<0.001) and the Cardiorenal cluster (HR 2.03, 95% CI 1.54–2.67; P<0.001). The Metabolic/obese cluster exhibited an intermediate risk profile (HR 1.11, 95% CI 0.83–1.48; P = 0.49), which was not significantly different from the reference group. Kaplan–Meier curves demonstrated clear separation between survival trajectories across phenotypes (log-rank P<0.001), with the Chagasic group exhibiting the steepest decline in event-free survival. (Figure 1)

A total of 1,539 heart failure hospitalisations occurred during follow-up. In Cox proportional hazards models, hospitalization risk was significantly elevated in all phenotypes relative to the Young group, with the highest risk in the Chagasic cluster (HR 3.97, 95% CI 3.26–4.83; P<0.001), followed by the Ischaemic High-Risk cluster (HR 2.77, 95% CI 2.36–3.25; P<0.001), the Cardiorenal cluster (HR 1.86,

95% CI 1.56–2.21; P<0.001), and the Metabolic/Obese cluster (HR 1.25, 95% CI 1.06–1.48; P=0.009). Consistent with the mortality analysis, Kaplan–Meier estimates showed a proportional and stepwise separation of hospitalisation curves across clusters (log-rank P < 0.001). (Figure 2)

When analysed using Fine–Gray subdistribution hazards to account for death as a competing event, differences in the cumulative incidence of first HF hospitalisation were attenuated and no longer statistically significant (Gray's test P=0.12). Compared with the Young cluster, the Chagasic phenotype had a subdistribution hazard ratio (SHR) of 1.27 (95% CI 0.83–1.94; P=0.28), while the Ischaemic High-Risk cluster showed an SHR of 1.17 (95% CI 0.82–1.67; P=0.38), the Cardiorenal cluster an SHR of 1.03 (95% CI 0.68–1.57; P=0.88), and the Metabolic/obese cluster an SHR of 0.87 (95% CI 0.58–1.29; P=0.48). Multivariable models further identified higher estimated glomerular filtration rate (eGFR) (per 10 mL/min; SHR 0.92, 95% CI 0.90–0.95; P<0.001) and higher systolic blood pressure (per 10 mm Hg; SHR 0.94, 95% CI 0.91–0.97; P<0.001) as independent protective factors, whereas increasing age was paradoxically associated with lower hospitalization risk (SHR 0.98 per year; P<0.001), likely reflecting survivor or competing-risk bias. Neither LVEF, BNP, nor baseline triple therapy independently predicted hospitalisation risk after adjustment for other variables. Sensitivity analyses using Fine–Gray models are presented in the Supplementary Material.

During the study period, 2,591 emergency department (ED) visits for acute decompensated HF were recorded. Event rates were significantly higher across all phenotypes compared with the Young reference. The Chagasic cluster had the most significant risk (HR 2.81, 95% CI 2.40–3.30; P<0.001), followed by the Ischaemic High-Risk cluster (HR 2.18, 95% CI 1.95–2.44; P<0.001), the Cardiorenal cluster (HR 1.70, 95% CI 1.51–1.91; P<0.001), and the Metabolic/Obese cluster (HR 1.35, 95% CI 1.18–1.54; P<0.001). Cumulative event-time analyses confirmed parallel gradients of risk across all endpoints (log-rank P < 0.001), reflecting consistent patterns of clinical severity and healthcare utilisation among the phenotypes. (Figure 3)

Discussion

In this large real-world cohort of patients with chronic HFrEF, we identified five clinically recognisable machine learning-derived phenotypes with distinct profiles and consistently different risks of all-cause mortality, heart failure hospitalisation and ED visits. Although the Ischaemic high-risk and Chagasic phenotypes showed higher cumulative incidence of first HF hospitalisation and mortality, these differences were attenuated after multivariable adjustment in Fine–Gray competing-risk models. This pattern suggests that the excess risk in these subgroups may be largely mediated by baseline clinical severity and comorbid burden, rather than by intrinsic phenotypic features alone. Conversely, the Metabolic/obese and Young clusters maintained lower event rates, consistent with a more compensated or early-stage HF profile.

The young phenotype had the most favourable outcomes. These patients were significantly younger, with relatively preserved renal function, higher rates of guideline-directed medical therapy, and the lowest hazard ratios across all outcomes. In contrast, the chagasic and ischaemic high-risk clusters had markedly elevated mortality and healthcare utilisation. [12] Notably, renal function and systolic blood pressure emerged as independent predictors of hospitalisation and mortality, renal impairment emerged as a critical determinant of adverse outcomes consistent with prior studies and underscoring the interplay between renal–cardiac coupling and hemodynamic stability in advanced HF.[13] These findings reinforce the importance of early identification and aggressive management of renal dysfunction in HFrEF patients; also, renal protective therapies may be a target priority in these clusters.

The chagasic phenotype, characterised by severe disease manifestations, including the lowest systolic blood pressure, highest BNP levels, and widespread myocardial fibrosis, further highlights the unique challenges faced by patients with Chagas cardiomyopathy. [14-15] The poor adherence to triple therapy observed in this cluster likely contributes to its adverse outcomes and may reflect barriers to care, including socioeconomic or healthcare access challenges.[16-18] These findings suggest that targeted strategies to

improve medication adherence and optimise care delivery are crucial for enhancing outcomes in this high-risk group, particularly in lower-resource settings. [19-20]

The metabolic/obese phenotype exhibited the highest prevalence of obesity, diabetes, and hypertension—features commonly associated with adverse cardiometabolic remodelling. Despite this high comorbidity burden, these patients had intermediate clinical outcomes, possibly due to preserved renal function and relatively maintained LVEF. Recent data suggest that obesity may be associated with a paradoxical survival benefit in HFrEF, although differences in muscle mass, disease stage, and inflammation likely confound this effect. Notably, the emergence of novel antiobesity therapies—particularly GLP-1 receptor agonists and dual GLP-1/GIP agonists—has shown promise in reducing cardiovascular risk in patients with obesity and diabetes. Their potential role in HFrEF patients with cardiometabolic phenotypes warrants further investigation, especially considering the rising prevalence of obesity-driven heart failure. Incorporating these agents into future phenotype-guided clinical trials may offer new therapeutic opportunities for this subgroup. [20]

Overall, our findings reinforce the concept that HFrEF is not a homogeneous condition. Data-driven phenotyping revealed reproducible gradients of risk that extend beyond LVEF and NYHA class, supporting the potential of personalised treatment approaches. These results should stimulate the design of phenotype-guided clinical trials, particularly targeting high-risk groups such as patients with renal impairment or Chagas cardiomyopathy.

Limitations

This study has limitations that warrant consideration. First, its retrospective cohort design restricts causal inference between phenotypic characteristics and clinical outcomes. [23] Although multiple imputation and mixed-data factor analysis were used to minimise bias, missing data and model assumptions remain potential sources of uncertainty. K-means clustering with Euclidean distance may not capture nonlinear relationships, underscoring the need for validation using alternative algorithms. Moreover, post hoc

stratification by outcome could introduce interpretative bias when assigning clinical labels. [24] Despite these constraints, the large sample size, diversity, and real-world setting enhance external applicability.[25] Internal Bootstrap validation supports the internal robustness. [26] Important determinants, such as medication adherence, socioeconomic status, and access to advanced therapies, were unavailable and may confound phenotypic risk stratification. Lastly, Kaplan–Meier estimates became sparse at later follow-up, suggesting cautious interpretation beyond five years.[27] Prospective multicenter studies should confirm these findings and refine phenotype-guided management in HFrEF

Conclusion

Machine-learning phenotyping of real-world HFrEF identified five clinically distinct subgroups with divergent risks for death and healthcare utilisation. Chagasic, ischaemic and cardiorenal phenotypes highlight the need for phenotype-tailored approaches and equitable resource allocation. Our results suggest that phenotype-aware approaches could be explored in future prospective studies and trials; however, external validation and implementation research are required before their routine adoption.

Table 1 - Baseline characteristics by phenotype

Variable	Overall (n = 10,376)	Ischaemic High-Risk N = 2,881	Chagasic N = 544	Cardiorenal N = 1,677	Metabolic/Obese N = 2,237	Young N = 3,037	P Value[2]
Age, years[1]	56 (47–63)	58 (52–65)	56 (46–63)	63 (56–70)	57 (51–64)	47 (37–56)	<0.001
Male, %[3]	64	66	59	54	69	65	<0.001
LVEF, %[1]	34 (26–47)	27 (21–33)	29 (22–41)	43 (31–58)	38 (30–49)	40 (30–55)	<0.001
eGFR, mL/min/1.73 m² [1]	64 (47–83)	55 (40–69)	53 (37–70)	45 (29–61)	64 (50–80)	85 (72–99)	<0.001
SBP, mm Hg[1]	120 (110–130)	110 (100–120)	110 (100–120)	120 (110–140)	124 (120–140)	120 (110–126)	<0.001
BMI, kg/m² [1]	26.2 (23.4–29.6)	24.5 (22.2–27.1)	23.9 (21.8–26.7)	26.6 (24.0–30.1)	30.1 (26.1–32.3)	26.1 (23.3–29.3)	<0.001
Diabetes mellitus, %[3]	29	21	17	33	75	4	<0.001
Hypertension, % [3]	36	15	22	60	71	20	<0.001
Atrial fibrillation, % [3]	30	37	48	27	31	21	<0.001
Chagas etiology, % [3]	5.6	0.6	99	0.9	0.4	0.2	<0.001
BNP, pg/mL [1]	176 (57–326)	244 (158–573)	228 (176–489)	176 (68–351)	116 (41–232)	76 (27–176)	<0.001
Triple therapy (ARNI/ACEI + MRA + β-blocker), % [3]	53	72	58	12	73	49	<0.001
All-cause mortality, %[3]	6.7	9.4	24.0	6.3	4.4	3.0	<0.001
HF hospitalization, %[3]	15	20	31	15	13	8	<0.001
ED visit, %[3]	25	31	42	26	26	16	<0.001

[1]Median (IQR)
[2]Kruskal-Wallis rank sum test; Pearson's Chi-squared test
[3]Proportions (%)

Subtitles: Neprilysin receptor antagonists (ARNI); angiotensin II receptor blocker (ARB); Mineralocorticoid receptor antagonist (MRA); Left ventricle ejection fraction (LVEF); Angiotensin converting enzyme inhibitor (ACEI); ACEI/ARNI/ARB + beta-blockers + mineralocorticoid receptor antagonist (MRA - triple therapy; systolic blood pressure (SBP); estimated glomerular filtration rate (eGFR), brain natriuretic peptides (BNP).

5 RESULTS

This integrated analysis evaluated three complementary dimensions of heart failure (HF) care in a large real-world cohort: (1) the relationship between guideline-directed medical therapy (GDMT) dose intensity and the risk of hyperkalemia and mortality in patients with heart failure with reduced ejection fraction (HFrEF); (2) imaging-based phenotyping using cardiac magnetic resonance (CMR) and its association with structural recovery potential and mortality; and (3) the prognostic utility of B-type natriuretic peptide (BNP) across clinically relevant subgroups. Together, these findings underscore the multidimensional complexity of HF risk stratification and therapeutic safety, particularly in resource-limited environments where treatment optimization must balance efficacy, tolerability, and access.

5.1.1 Guideline-Directed Medical Therapy and Risk of Hyperkalemia

Study and population

A retrospective cohort of 1,023 patients with heart failure with reduced ejection fraction (HFrEF, LVEF ≤40%) and a history of mild hyperkalemia was analyzed. Guideline-directed medical therapy (GDMT) use—including ACE inhibitors (ACEIs), angiotensin receptor blockers (ARBs), mineralocorticoid receptor antagonists (MRAs), and beta-blockers (BBs)—was recorded and standardized as a percentage of the guideline-recommended target dose. Exposure scenarios were categorized as baseline (MIN), maximum achieved (MAX), and within-patient uptitration (Δ = MAX – MIN). Time-to-event analyses were performed using inverse probability of treatment weighting (IPTW) and Cox proportional hazards models, with log-time interaction terms for non-proportional hazards and Fine–Gray competing risk models for non-fatal outcomes.

Hyperkalemia Outcomes

Moderate hyperkalemia (K^+ ≥5.5 mmol/L) occurred in 30.3% of the cohort within 730 days. Risk patterns varied markedly by class and exposure:

- ACEI/ARB uptitration (Δ) was associated with a significantly increased risk of hyperkalemia: ACEI Δ aHR 1.34 (95% CI 1.07–1.67; p=0.011), ARB Δ aHR 1.39 (95% CI 1.07–1.75; p=0.009).
- Baseline beta-blocker exposure was protective (MIN aHR 0.85; 95% CI 0.74–0.97; p=0.015), with time-varying effects indicating increasing protection over follow-up.
- MRA and loop diuretic uptitration showed weak or neutral associations with hyperkalemia.
- Time-varying hazard models confirmed these trends, with beta-blocker effects becoming more protective over time, while ACEI/ARB exposure remained either neutral or detrimental.

Mortality Outcomes

In survival models:

- Higher baseline and achieved doses of beta-blockers and ACEIs were associated with significantly lower early mortality (BB: HR 0.80–0.82; ACEI: HR 0.63–0.73).
- However, time-varying models revealed attenuation of benefit over time, suggesting that early gains were partially offset by evolving risk or changes in therapy.
- MRAs were neutral across models.

- CKD, baseline hyperkalemia, atrial fibrillation, and Chagas disease were independent predictors of higher mortality.
- Sensitivity analyses (including trimmed weights, competing risks, and alternate potassium thresholds) did not materially change conclusions.

Dose-Response Associations with Mortality

Mortality analyses demonstrated time-dependent protective associations for ACEIs and beta-blockers. Higher baseline and achieved doses of beta-blockers were associated with early survival benefit (HRs 0.82 and 0.80, respectively), though both effects attenuated over time. ACEIs similarly conferred early protection (baseline HR 0.63; MAX HR 0.73), with significant time interactions. Conversely, the MRA dose showed no significant early or time-varying effects on mortality. Uptitration of ACEIs or ARBs (Δ) was not associated with independent survival benefit and, in some cases, correlated with increased risk, reflecting possible confounding by underlying clinical instability or hyperkalemia-induced treatment adjustments.

Baseline potassium, impaired renal function, atrial fibrillation, and Chagas disease emerged as independent predictors of increased mortality. These findings reinforce the need for individualized treatment strategies that account for electrolyte vulnerability and comorbidity burden when guiding GDMT titration.

5.1.2 Imaging-Derived Cardiomyopathy Phenotypes and Reverse Remodeling Potential

Study Design and Methods

Using data from 2,367 patients who underwent clinical cardiac magnetic resonance (CMR), a natural language processing (NLP) algorithm was developed to extract structured imaging data—including LV/RV volumes, ejection fractions, and late gadolinium enhancement (LGE)—from free-text reports. Unsupervised k-means clustering on standardized imaging variables identified structural phenotypes. A continuous reverse remodeling (RR) score was computed via machine learning models (Random Forest), trained to predict theoretical recovery capacity based on imaging features. Outcomes included all-cause mortality and reverse remodeling.

Phenotypic Clustering and Structural Heterogeneity

Three distinct clusters emerged:

- **Cluster 1**: Mild remodeling, preserved systolic function, minimal fibrosis.
- **Cluster 2**: Ischemic phenotype with large LV volumes, biventricular dysfunction, and high transmural LGE burden.
- **Cluster 3**: Markedly remodeled non-ischemic pattern with large LVESVi, low LVEF, but intermediate scar burden.

Principal component analysis (PCA) confirmed clear phenotypic separation. Cluster 2 was associated with worse renal function, higher BNP levels, and lower systolic blood pressure.

Reverse Remodeling Potential and Prognostic Impact

The continuous RR score (range: 0 to 1) differed significantly across phenotypes:

- Highest in Cluster 3 (mean RR = 0.726), intermediate in Cluster 1 (0.584), and lowest in ischemic Cluster 2 (0.279).
- RR score was a strong predictor of mortality (HR per unit increase: 0.03; 95% CI 0.004–0.18; $p < 0.001$), outperforming LVEF.
- In fully adjusted Cox models, LVEF was not independently predictive ($p = 0.87$), while RR remained robust.
- Scar burden (% LGE) was a stronger predictor than LGE pattern, challenging traditional morphology-based risk schemes.
- Random Forest classifiers showed excellent accuracy (error rate: 5.28%), with high RR patients never misclassified as low RR.

5.1.3 BNP and special populations

Study Population

A total of 10,376 ambulatory patients with heart failure were included in the analysis. The median age of the cohort was 57 years (interquartile range [IQR] 47–64), and 63% were male. During follow-up, 2,538 patients (24.5%) died, while 7,838 were alive or censored.

Compared with survivors, patients who died were older and exhibited a higher burden of comorbidities, including chronic kidney disease (CKD) and atrial fibrillation. Renal function was significantly worse among deceased patients, with lower estimated glomerular

filtration rate (eGFR). Median baseline BNP levels were substantially higher in patients who died compared with those who survived (191 pg/mL [IQR 144–473] vs. 138 pg/mL [IQR 46–277]; $p < 0.001$)..

Discriminatory Performance of BNP for Mortality

Overall, baseline BNP demonstrated moderate discriminatory performance for all-cause mortality, with an area under the receiver operating characteristic (ROC) curve (AUC) of 0.65. The optimal BNP cutoff derived from the Youden Index was 175.5 pg/mL, yielding a sensitivity of 71% and a specificity of 56%.

When stratified by clinical subgroups, the discriminatory performance of BNP varied modestly. In patients with CKD, the AUC was 0.61, whereas in those without CKD it increased to 0.66. Among obese patients, BNP exhibited a lower optimal cutoff (114.5 pg/mL), while maintaining a similar AUC (0.64) compared with non-obese individuals. Discrimination was reduced in patients aged ≥65 years and in those with atrial fibrillation (AUCs of 0.61 and 0.62, respectively). No meaningful differences in AUC were observed between sexes.

Survival Analyses According to BNP Levels

Kaplan–Meier survival analyses demonstrated significantly lower survival among patients with BNP levels above the subgroup-specific cutoff compared with those below the cutoff across all examined strata (log-rank $p < 0.001$ for all comparisons). This association was consistent in patients with and without CKD, in obese and non-obese individuals, in those with atrial fibrillation, and across age groups.

The magnitude of survival separation was particularly pronounced in patients with CKD, atrial fibrillation, and older age, underscoring the robust prognostic value of BNP even in subgroups traditionally considered challenging for biomarker interpretation.

Association Between BNP and Mortality in Cox Regression Models

In unadjusted Cox proportional hazards models, higher BNP levels were associated with an increased risk of all-cause mortality. After adjustment for age, sex, chronic kidney disease, obesity, and atrial fibrillation, elevated BNP remained independently associated with mortality.

When BNP was modeled as a continuous variable using restricted cubic splines, a non-linear association with mortality risk was observed. Mortality risk increased progressively with rising BNP concentrations, with a noticeable inflection beginning at approximately 100 pg/mL, followed by a steeper increase at higher levels.

6 DISCUSSION

Principal Findings

This thesis presents a comprehensive, multidimensional evaluation of chronic heart failure with reduced ejection fraction (HFrEF), using clinical, imaging, biomarker, and phenotypic data from over 10,000 patients managed at a tertiary HF center in Brazil. Across five original investigations, we identified novel insights that reinforce the central theme of this work: HFrEF is a biologically heterogeneous syndrome requiring phenotype-guided risk stratification and personalized management**.**

Our findings encompass:

- The dynamic impact of guideline-directed medical therapy (GDMT) exposure on potassium disturbances and mortality;
- A novel, continuous CMR-derived reverse remodeling (RR) probability score predictive of long-term survival;
- The context-specific prognostic value of BNP across clinically relevant subgroups;
- The identification of five real-world phenotypes of HFrEF using unsupervised machine learning;
- The interaction between potassium disorders, HF etiology (particularly Chagas), and competing modes of death.

Taken together, these studies converge to demonstrate that risk in HFrEF is multifactorial, trajectory-dependent, and phenotype-specific**.** Traditional metrics such as LVEF and NYHA class are insufficient to fully capture this complexity.

6.1.1 Machine Learning Phenotypes and Structural Heterogeneity

Our unsupervised clustering study (n = 10,376) revealed five clinically distinct HFrEF phenotypes with divergent risk profiles: Young**,** Chagasic**,** Ischaemic High-Risk**,** Cardiorenal**,** and Metabolic/Obese**.**

These phenotypes were built using a data-driven framework that combined clinical, laboratory, therapy, and comorbidity data using factorial analysis of mixed data and k-means clustering. Each cluster demonstrated distinct trajectories:

- The Chagasic phenotype had the highest mortality (24%) and HF hospitalization (31%), with severe biventricular dysfunction, low systolic blood pressure, and poor GDMT uptake.
- The Ischaemic High-Risk group featured the second-highest mortality and extensive coronary disease despite good therapy use, suggesting refractory disease biology.
- The Cardiorenal cluster displayed high AF burden and the worst renal function, aligning with prior evidence of renal dysfunction as a dominant prognostic driver in HFrEF.
- The Metabolic/Obese cluster showed intermediate outcomes, despite high obesity and diabetes prevalence, possibly reflecting preserved perfusion and "obesity paradox" phenomena.
- The Young group, with preserved renal function and minimal comorbidities, had the lowest risk, consistent with early-stage or genetically driven disease.

These phenotypes aligned with biological mechanisms, and their consistent risk gradients across mortality, hospitalization, and emergency utilization support their clinical utility in stratification and trial design.

6.1.2 3. Drug Exposure, Hyperkalemia, and RAASi Risk Profiles

We showed that RAAS inhibitor uptitration — particularly of ACEIs and ARBs — was independently associated with an increased risk of hyperkalemia, even after adjustment for baseline eGFR and potassium levels. Notably, the timing and intensity of titration was more predictive of risk than the absolute dose itself. Beta-blockers, conversely, conferred a time-varying protective effect on mortality and hyperkalemia, likely due to modulation of sympathetic tone and renal perfusion.

These results resonate with those of Lima et al., who found that potassium disorders were strongly associated with mode of death, and that hypokalemia was particularly lethal in Chagas cardiomyopathy, while hyperkalemia was most associated with renal disease and RAASi exposure.

Collectively, these findings suggest that:

- RAASi titration should not be viewed as universally safe**,** especially in the Chagasic and Cardiorenal phenotypes**,** where underlying physiology and access to monitoring may amplify risk.
- Therapy monitoring must be adapted not only to renal function, but also to underlying HF phenotype and etiology**.**

6.1.3 Imaging-Based Risk Stratification: Beyond LVEF

Using NLP to extract and transform unstructured CMR reports, we developed the reverse remodeling probability (RRp) score — a multidimensional continuous measure integrating scar burden, chamber volumes, wall motion, and RV function. RRp outperformed LVEF in

predicting mortality and identified patients with structural potential for recovery, particularly in non-ischemic phenotypes.

Notably, fibrosis burden , not just pattern , was the dominant predictor of RRp and outcomes. This reinforces prior evidence that scar extent limits remodeling**,** regardless of ischemic etiology. Additionally, the inclusion of RV dysfunction as a key feature reflects the known biventricular nature of advanced HFrEF**.**

This tool can support real-world risk assessment**,** especially where manual quantification is unavailable, and it aligns well with phenotype-guided care strategies emerging from our clustering work.

6.1.4 BNP Interpretation: A Case for Precision Thresholds

Our BNP analysis across 6,000+ patients demonstrated that:

- BNP remains a robust predictor of mortality;
- But optimal cutoffs vary widely by subgroup**:** lower in obese patients, higher in CKD and AF;
- Spline models showed an exponential increase in mortality risk above 100 pg/mL, but the slope differed by comorbidity.

This underscores that a fixed BNP threshold is insufficient in real-world settings, and that biomarker interpretation must be contextualized by patient phenotype**.**

6.1.5 Clinical Implications and Phenotype-Aware Medicine

These findings point to immediate clinical applications:

- Chagasic patients: High structural risk, poor therapy adherence, and increased potassium sensitivity demand individualized surveillance and targeted interventions, especially in low-resource environments.
- Cardiorenal patients: Require cautious uptitration, renal-protective strategies, and close potassium monitoring.
- Young/Obese phenotypes: May tolerate aggressive therapy but benefit from early cardiometabolic intervention (e.g., GLP-1 agonists).
- Reverse remodeling score: Could guide decision-making regarding CRT, transplantation, or intensified therapy.
- BNP thresholds: Should be adapted to comorbid states within EHR-based alerts.

This framework supports a shift toward precision medicine in HF, moving beyond LVEF-centric and class-based decisions toward integrated, phenotype-aware approaches**.**

6.1.6 Limitations

These studies share some limitations:

- Observational design limits causal inference despite use of IPTW and sensitivity analyses;
- NLP model performance depends on local reporting structure;
- Phenotypes were derived in a single-center cohort with high Chagas prevalence, which may limit external generalizability;
- BNP analyses lacked serial measurements or NT-proBNP data.

Nonetheless, large sample size, detailed phenotyping, and internal validation strengthen the robustness and relevance of these findings.

7 CONCLUSIONS

This doctoral thesis contributes to the evolving paradigm of heart failure management by offering a multidimensional, phenotype-aware framework for understanding prognosis and therapeutic response in patients with heart failure with reduced ejection fraction (HFrEF). Through the integration of real-world clinical data, advanced imaging, laboratory biomarkers, and machine learning–derived phenotypes, we demonstrate that the trajectory and outcomes of HFrEF are shaped not solely by left ventricular ejection fraction or guideline adherence, but by a complex interplay of structural capacity, comorbid burden, treatment tolerability, and biochemical disturbance.

We identified five clinically relevant HFrEF phenotypes using unsupervised clustering, each with distinct comorbidity profiles, mortality risk, and patterns of therapy utilization. These phenotypes, especially the Chagasic and cardiorenal subgroups, illustrated the unequal distribution of risk and underscored the need for individualized therapeutic strategies beyond conventional metrics.

Our analyses further revealed that uptitration of renin–angiotensin system inhibitors, while beneficial for many, carries a time-dependent risk of hyperkalemia and mortality that varies by underlying phenotype and renal function. Beta-blockers were protective, particularly when initiated early and maintained at higher doses. These findings support a more cautious and personalized approach to drug escalation in vulnerable subgroups.

In imaging, we proposed and validated a continuous reverse remodeling probability (RRp) score derived from routine CMR reports, which robustly predicted long-term survival and captured recovery potential beyond ejection fraction. This tool represents a scalable

innovation for real-world prognostic stratification, particularly in non-ischemic or biventricular disease.

Finally, we showed that the prognostic performance of BNP, a cornerstone biomarker in HF, is modulated by clinical context. Optimal thresholds varied significantly across subgroups such as obesity, CKD, and atrial fibrillation, advocating for the use of individualized cutoffs and nonlinear modeling in both clinical practice and research.

Together, these findings support the implementation of precision strategies in HFrEF care. Such strategies include phenotype-guided titration protocols, imaging-informed risk stratification, biomarker contextualization, and real-time integration of clinical data for decision support. By embracing the biological and therapeutic heterogeneity of heart failure, we can improve risk prediction, optimize treatment allocation, and ultimately enhance outcomes for diverse patient populations.

Future work should focus on validating these findings in multicenter cohorts, testing phenotype-specific interventions in clinical trials, and embedding these tools into electronic health systems to bridge the gap between data-rich research and day-to-day care. As heart failure remains one of the leading causes of morbidity and mortality globally, advancing toward a more personalized, data-driven, and equitable model of care is not only feasible, it is imperative.

REFERENCES*

* According to Vancouver Style.

REFERENCES-

1. Roth GA, Shahraz S, Forouzanfar MH, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol*. 2017;70(1):1–25.
2. Shahim B, Yu G, Agarwal S, et al. Global, regional, and national burden of heart failure, 1990–2020. *JAMA Cardiol*. 2023;8(4):370–380.
3. Chen C, Zhao D, Du X, et al. Global trends in heart failure prevalence and risk factors from 1990 to 2025. *Nature Reviews Cardiology*. 2025;22:205–214.
4. Pianca EG, Heidemann AI, Santos ÂBS, et al. Prevalence and sex differences in stages of heart failure: insights from the ELSA-Brasil cohort. *Eur Heart J*. 2024;45(Suppl_1). https://doi.org/10.1093/eurheartj/ehae666.892
5. Arruda G, Souza Neto JD, Melo DM, et al. Trends in heart failure-related hospitalizations and outcomes in Brazil: a national analysis. *Arq Bras Cardiol*. 2022;119(3):450–460.
6. Fernandes F, Bocchi EA, Oliveira WA, et al. Temporal trends in mortality and length of stay for heart failure in Brazil. *Rev Soc Bras Med Trop*. 2020;53:e20190237.
7. Gargiulo L, Suárez Fernández C, Rossignol P, et al. Comorbidities and multimorbidity in heart failure: clinical challenges and opportunities. *Eur J Heart Fail*. 2022;24(5):794–803.
8. Ergatoudes C, Schaufelberger M, Andersson B, et al. Non-cardiac comorbidities and survival in heart failure with reduced vs. preserved ejection fraction: data from the Swedish Heart Failure Registry. *Eur J Heart Fail*. 2019;21(9):1208–1217.
9. Christiansen MN, Kober L, Weeke P, et al. Impact of age and comorbidity burden on mortality in heart failure: a nationwide cohort study. *Eur J Heart Fail*. 2020;22(5):891–899.
10. Galli E, Stolfo D, Merlo M, et al. Unsupervised machine learning for phenotyping in HFpEF: mechanisms and prognosis. *Eur J Heart Fail*. 2021;23(6):887–899.
11. Jones NR, Roalfe AK, Adoki I, et al. Cluster-based phenotyping in heart failure with preserved ejection fraction. *Heart*. 2021;107(9):709–716.
12. Kintscher U, Edelmann F. The non-steroidal mineralocorticoid receptor antagonist finerenone and heart failure with preserved ejection fraction. *Cardiovasc Diabetol*. 2023;22:30. doi:10.1186/s12933-023-01899-0
13. Gouda P, Gomaa R, Omran J, et al. Clinical phenotypes and outcomes in HFrEF: insights from machine learning clustering. *ESC Heart Fail*. 2022;9(4):2486–2495.
14. Balletti S, Esposito M, De Filippo O, et al. HbA1c and inflammatory signatures in heart failure subtypes. *Int J Cardiol*. 2023;380:57–65.
15. Brown P, Shah M, Dweck M, et al. Cardiovascular magnetic resonance in HFmrEF: insights into an intermediate phenotype. *JACC Cardiovasc Imaging*. 2022;15(4):623–634.
16. Yogasundaram H, Mamdani M, Park J, et al. Cardiorenal syndrome: mechanisms and therapeutic targets. *J Am Coll Cardiol*. 2019;73(3):423–438.
17. Chávez-Iñiguez JS, Venegas-Gómez A, Vázquez-Rangel A, et al. Diagnosis and management of cardiorenal syndrome: clinical practice update. *Kidney Int Rep*. 2022;7(4):861–871.
18. Khandait D, Tumlin J, Gupta S. Electrolyte abnormalities in cardiorenal syndrome: diagnostic and prognostic considerations. *Heart Fail Rev*. 2025.