IS 790: Causal Artificial Intelligence

Spring 2024

Professor : Md Osman Gani

Final Project Report

on

Heart Failure Analysis

Submitted By :

Catherin Revu

Parthiv Gandhi

Shanmukhi Gundu

Sathwik Reddy Gangannagari

Information Systems Department

University of Maryland, Baltimore County

Table of Contents

1. Introduction

Heart failure, a condition marked by the heart's inability to pump blood efficiently, is a critical global health issue affecting millions worldwide. This project centers on analyzing clinical records from 299 patients with heart failure, aiming to decipher the complex interplay of various factors leading to this condition. Heart failure not only compromises individual well-being but also imposes significant burdens on healthcare systems due to its chronic nature and the extensive care required for management. The complexity arises from the multifaceted nature of heart failure, influenced by a spectrum of risk factors including age, lifestyle choices, and pre-existing conditions like diabetes and high blood pressure. Our in-depth analysis seeks to explore these variables, uncovering causal relationships and identifying patterns that could predict heart failure outcomes more accurately.

Importance of the Problem

The significance of addressing heart failure lies in its profound impact on public health. As the leading cause of hospitalization among adults over the age of 65, heart failure represents a pivotal challenge in cardiovascular medicine. The increasing prevalence of heart failure, coupled with high mortality rates, underscores the urgent need for advanced predictive models and effective treatment strategies. By leveraging causal artificial intelligence methodologies, this project aims to contribute to the early detection and personalized management of heart failure, potentially reducing its incidence and improving patient outcomes.

Visualizing the Problem

To explain the heart failure issue comprehensively, several key figures are proposed :

Global Prevalence and Mortality Trends: A line graph illustrating the rise in heart failure cases and associated mortality over recent years.
Risk Factor Distribution: Histograms or pie charts showing the prevalence of critical risk factors such as diabetes, hypertension, and obesity within the patient cohort.
Correlation Heatmap: A heatmap detailing the correlation between various clinical features and their relationship with heart failure outcomes, highlighting potential areas for causal investigation.

Challenges in Heart Failure Analysis

The analysis of heart failure comes with challenges, primarily due to the heterogeneity in patient symptoms, the overlap with other conditions, and the intricate web of contributing factors.

Diagnosing and predicting heart failure outcomes is complicated by the variability in how the condition manifests across individuals, necessitating a nuanced approach that considers a wide array of variables. Moreover, the reliance on high-quality, comprehensive clinical data poses significant challenges in ensuring the accuracy and applicability of predictive models.

Approach and Objectives

Our project employs a structured approach to dissect the complexities of heart failure through the lens of causal AI. By constructing and analyzing Structural Causal Models (SCMs) and applying principles of Do-Calculus, we aim to:

Identify Causal Relationships: Determine the direct and indirect factors contributing to heart failure, offering insights into how various conditions and lifestyle choices influence its development.

Develop Predictive Models: Create models capable of predicting heart failure outcomes with high accuracy, utilizing identified causal relationships to enhance prediction quality.

Contributions of the Project

This project is poised to make significant contributions to the field of heart failure research and management. By elucidating the causal pathways leading to heart failure, we can inform the development of targeted interventions and prevention strategies. Furthermore, the predictive models derived from this study have the potential to revolutionize patient care, enabling personalized treatment plans that could significantly improve outcomes for individuals at risk of heart failure. Ultimately, this project represents a step forward in the integration of causal AI into healthcare, paving the way for more informed, effective approaches to combating heart failure.

2. Prior Work

A brief summary prior and related works in the area:

Prior research has explored various machine learning approaches to predict heart failure based on clinical data. Some key studies include:

Study by (Mahmud et al., 2023) proposed a machine learning metamodel that combined Random Forest Classifier, Gaussian Naive Bayes, Decision Tree, and k-Nearest Neighbor models to predict heart failure with an accuracy of 87%.

Wang et al. (Moreno-Sánchez, 2021) conducted a comparative study of 18 popular machine learning models for heart failure prediction, finding that z-score normalization and the Synthetic Minority Oversampling Technique (SMOTE) improved model performance.

Researchers have also explored the use of deep learning (Wang, 2021) and explainable AI techniques to enhance the interpretability and accuracy of heart failure prediction models.

Zhang et al. (2023) leveraged bidirectional Mendelian randomization to illuminate the genetic causal pathways between atrial fibrillation and heart failure. Their study, affirming a reciprocal causal relationship, enriches the field with a genetic perspective that may inform personalized treatment strategies and early intervention in heart disease management.

Critical review of prior work:

The prior studies demonstrate the potential of machine learning to improve heart failure prediction compared to traditional risk assessment methods. However, several limitations exist:

Many studies have used relatively small datasets (e.g., 300 records) (Mahmud et al., 2023), which may limit the generalizability of the models.

There is a lack of consensus on the most relevant clinical features for heart failure prediction, with different studies using varying feature sets (Moreno-Sánchez, 2021)(Wang, 2021).

Interpretability of the prediction models is often limited, making it difficult to understand the key drivers of heart failure risk (Wang, 2021).

Few studies have explicitly addressed the class imbalance problem inherent in heart failure datasets, where the number of patients who experience a death event is much lower than those who do not (Moreno-Sánchez, 2021).

Prior work that we will be building on:

This project will build upon the work of (Mahmud et al., 2023) by further exploring the use of ensemble machine learning models for heart failure prediction. Additionally, the project will leverage the insights from Wang et al. (Moreno-Sánchez, 2021) regarding the benefits of data normalization and oversampling techniques to address class imbalance.

Gaps in Current Research

1. **Lack of Comprehensive Causal Analysis**: Many existing studies on heart failure prediction, such as those by Mahmud et al. (2023) and Wang (2021), primarily focus on predictive modeling using various machine learning techniques. While these studies have advanced the prediction of heart failure, there is a notable gap in the comprehensive analysis of causal relationships between different clinical variables and heart failure outcomes.

2. **Insufficient Focus on Early Intervention**: The identification of early intervention points has not been the primary focus of much of the current research. Studies often concentrate on predicting heart failure or mortality risk without providing actionable insights into when and how interventions could be most effectively applied.
3. **Limited Use of Bayesian Networks in Heart Failure Research**: Despite the potential of Bayesian network models to understand complex interrelationships and conditional dependencies between clinical factors, their application in the context of heart failure prediction has been limited.

How do you expect to improve on your/others' prior work?:

Our project's integration of Structural Causal Models and Bayesian networks for analyzing heart disease data represents a novel approach that goes beyond traditional predictive modeling. Unlike the studies by Mahmud et al. (2023) and Wang (2021), which focus on machine learning for prediction, our project seeks to uncover the underlying causal mechanisms of heart failure, thereby offering potential for more targeted and effective interventions.

Additionally, by aiming to enhance diagnostic accuracy and facilitate early intervention, our project contributes to a more proactive healthcare approach, aligning with the needs for precision medicine and personalized treatment strategies.

Innovative Application of Causal AI: By leveraging Structural Causal Models and Bayesian network models, our project innovates in understanding the complex causal relationships that contribute to heart failure. This approach not only aids in predicting heart failure but also in identifying the underlying causes and effects, providing a deeper understanding of the disease dynamics.

Enhanced Early Decision-Making: Our focus on enabling earlier diagnosis and intervention decisions directly addresses the critical gap of timely prevention and treatment of heart failure. This early decision-making capability could significantly alter patient outcomes by preventing disease progression.

Identification of Early Intervention Points: By pinpointing critical areas where interventions could substantially lower heart failure risks, our project provides valuable insights that can guide healthcare professionals in prioritizing and implementing preventive measures.

Expanding Causal Inferences in Heart Failure Research: Our work aims to expand on the findings of Zhang et al. (2023), who used Mendelian randomization to give evidence of a bidirectional causal link between atrial fibrillation and heart failure. This will fill in gaps in the current literature. Although this study provides a novel viewpoint on the genetic connections between these disorders, our work aims to improve heart failure prediction models by using a comparable causal analysis methodology to a larger range of clinical characteristics. Our goal is

to better understand the complex etiology of heart failure by utilizing modern computational tools and combining a wider range of clinical data. This method is well-positioned to pinpoint new therapy targets and improve preventative techniques, meeting the demand for a more individualized approach to the management of heart failure.

3. Background

a. Theoretical Foundations

This project leverages causal inference principles to identify relationships that influence heart failure, going beyond mere correlations to understand causative mechanisms. Key components include:

- **Structural Causal Models (SCMs):** These models express how variables causally affect each other, encapsulating the system's causal mechanisms.
- **Do-Calculus:** This methodology, developed by Judea Pearl, is used for deriving causal relationships from observational data combined with experimental conditions, facilitating the evaluation of intervention effects.

b. Mathematical Definitions and Algorithmic Foundations

- **Causal Graphs:** These are directed graphs where nodes represent variables and edges denote direct causal influences. They provide a visual and mathematical way to represent causality.

Definition: If node X directly causes node Y, represented by $Y=f(X,U)$ it implies X influences Y with U as other contributing factors.

- **Bayesian Networks:** These probabilistic models use a directed acyclic graph (DAG) to represent conditional dependencies among variables, crucial for computing probabilities and inferring intervention effects.

Definition: The joint probability distribution is given by:

$$P(X1,X2,...,Xn)=\prod i=1nP(Xi|\text{Parents}(Xi))$$

- **PC and FCI Algorithms:** These algorithms identify causal structures from data. The PC algorithm uses conditional independence tests to refine a fully connected graph into a Partially Directed Acyclic Graph (PDAG), while FCI extends this to handle latent variables and biases, resulting in a Partial Ancestral Graph (PAG).

These methods underpin our approach to analyzing heart failure, allowing for detailed exploration and understanding of the causal dynamics influencing patient outcomes.

4. Data

Dataset Overview

Source and Context: The data originates from the Faisalabad Institute of Cardiology and Allied Hospital in Faisalabad, Punjab, Pakistan, covering the period of April to December 2015. It encompasses 299 patients who have previously experienced heart failure and were classified under NYHA classes III or IV, indicating severe heart impairment.

Demographics: The cohort consists of 105 women and 194 men, aged between 40 and 95 years, all of whom have left ventricular systolic dysfunction. This demographic spread provides a substantial basis for analyzing heart failure across different age groups and sexes in a high-risk population.

Clinical Features

The dataset encapsulates 13 clinical, body, and lifestyle features, meticulously recorded to aid in the exploration of heart failure dynamics. These include:

1. **Age**: Represents the patient's age in years. Age is a crucial factor in heart disease, with risk increasing as people age. The range in this dataset is 40 to 95 years, covering a broad spectrum of adulthood into senior years.
2. **Anaemia**: A binary indicator (Yes=1, No=0) for the presence of anaemia, defined as a decrease in red blood cells or hemoglobin. Anaemia can exacerbate the symptoms of heart failure by reducing oxygen delivery to tissues.
3. **Creatinine Phosphokinase (CPK)**: The level of the CPK enzyme in the blood, measured in mcg/L. CPK levels rise in response to muscle damage, including heart muscle damage, making this a valuable marker for cardiac events.
4. **Diabetes**: A binary indicator (Yes=1, No=0) of whether the patient has diabetes. Diabetes is a significant risk factor for developing heart disease, as it can lead to changes in blood vessels and the heart muscle itself.
5. **Ejection Fraction**: Measures the percentage of blood leaving the heart at each contraction, indicating how well the heart is pumping. Values range from 14% to 80%, with lower percentages indicating worse heart function.
6. **High Blood Pressure**: A binary indicator (Yes=1, No=0) for the presence of hypertension. High blood pressure puts additional strain on the heart, contributing to the risk of heart failure.
7. **Platelets**: The concentration of platelets in the blood, measured in kiloplatelets/mL. Platelets play a critical role in blood clotting, with abnormalities potentially indicating underlying health issues, including heart disease.

8. **Serum Creatinine**: Indicates the level of serum creatinine in the blood, measured in mg/dL. This is a marker of kidney function, with higher levels suggesting renal dysfunction, which is often associated with heart failure.
9. **Serum Sodium**: The level of serum sodium in the blood, measured in mEq/L. Abnormal sodium levels can be a sign of heart failure or other conditions affecting the heart and kidneys.
10. **Sex**: A binary indicator (Female=0, Male=1) representing the patient's sex. Sex differences in heart failure presentations and outcomes are an area of growing research interest.
11. **Smoking**: A binary indicator (Yes=1, No=0) of whether the patient smokes. Smoking is a well-known risk factor for heart disease, contributing to arterial damage and heightened risk of cardiac events.
12. **Time**: Follow-up period in days, ranging from 4 to 285 days. This variable captures the duration over which each patient's health status was monitored, providing context for the progression or management of heart failure.
13. **[Target] Death Event**: A binary indicator (Yes=1, No=0) of whether the patient died during the follow-up period. This outcome variable enables predictive modeling of mortality risk based on the preceding clinical features.

Insights and Implications

Clinical and Lifestyle Interactions

The incorporation of lifestyle (such as smoking status) and clinical (such as ejection fraction and serum creatinine) variables enables a comprehensive examination of the complex interactions between these variables and the consequences of heart failure. For example, elevated CPK levels suggest muscle deterioration, which may signify sudden episodes of heart failure; smoking, among other lifestyle variables, may contribute to cardiac diseases.

Risk Factor Identification

The dataset supports the identification of key risk factors for heart failure mortality. Variables like ejection fraction and serum creatinine levels, which reflect heart and kidney function, respectively, are particularly telling. Lower ejection fractions and higher serum creatinine levels have been associated with poorer outcomes.

Gender and Age Dynamics

With a mix of 105 women and 194 men, the dataset facilitates gender-specific analyses, potentially uncovering differences in heart failure progression and outcomes across sexes. Age, as a variable, allows for age-specific risk assessments, offering critical insights into how heart failure impacts various age groups.

Statistical Patterns

The provided statistical breakdown (e.g., the percentage of patients with anaemia, diabetes, or high blood pressure) enables initial observations about the prevalence of these conditions within the heart failure population and their potential correlations with mortality rates.

Initial data exploration:

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	time	DEATH_EVENT
count	299.000000	299.000000	299.000000	299.000000	299.000000	299.000000	299.000000	299.00000	299.000000	299.000000	299.00000	299.000000	299.00000
mean	60.833893	0.431438	581.839465	0.418060	38.083612	0.351171	263358.029264	1.39388	136.625418	0.648829	0.32107	130.260870	0.32107
std	11.894809	0.496107	970.287881	0.494067	11.834841	0.478136	97804.236869	1.03451	4.412477	0.478136	0.46767	77.614208	0.46767
min	40.000000	0.000000	23.000000	0.000000	14.000000	0.000000	25100.000000	0.50000	113.000000	0.000000	0.00000	4.000000	0.00000
25%	51.000000	0.000000	116.500000	0.000000	30.000000	0.000000	212500.000000	0.90000	134.000000	0.000000	0.00000	73.000000	0.00000
50%	60.000000	0.000000	250.000000	0.000000	38.000000	0.000000	262000.000000	1.10000	137.000000	1.000000	0.00000	115.000000	0.00000
75%	70.000000	1.000000	582.000000	1.000000	45.000000	1.000000	303500.000000	1.40000	140.000000	1.000000	1.00000	203.000000	1.00000
max	95.000000	1.000000	7861.000000	1.000000	80.000000	1.000000	850000.000000	9.40000	148.000000	1.000000	1.00000	285.000000	1.00000

Figure 1. Data

Distribution of Numerical Features:

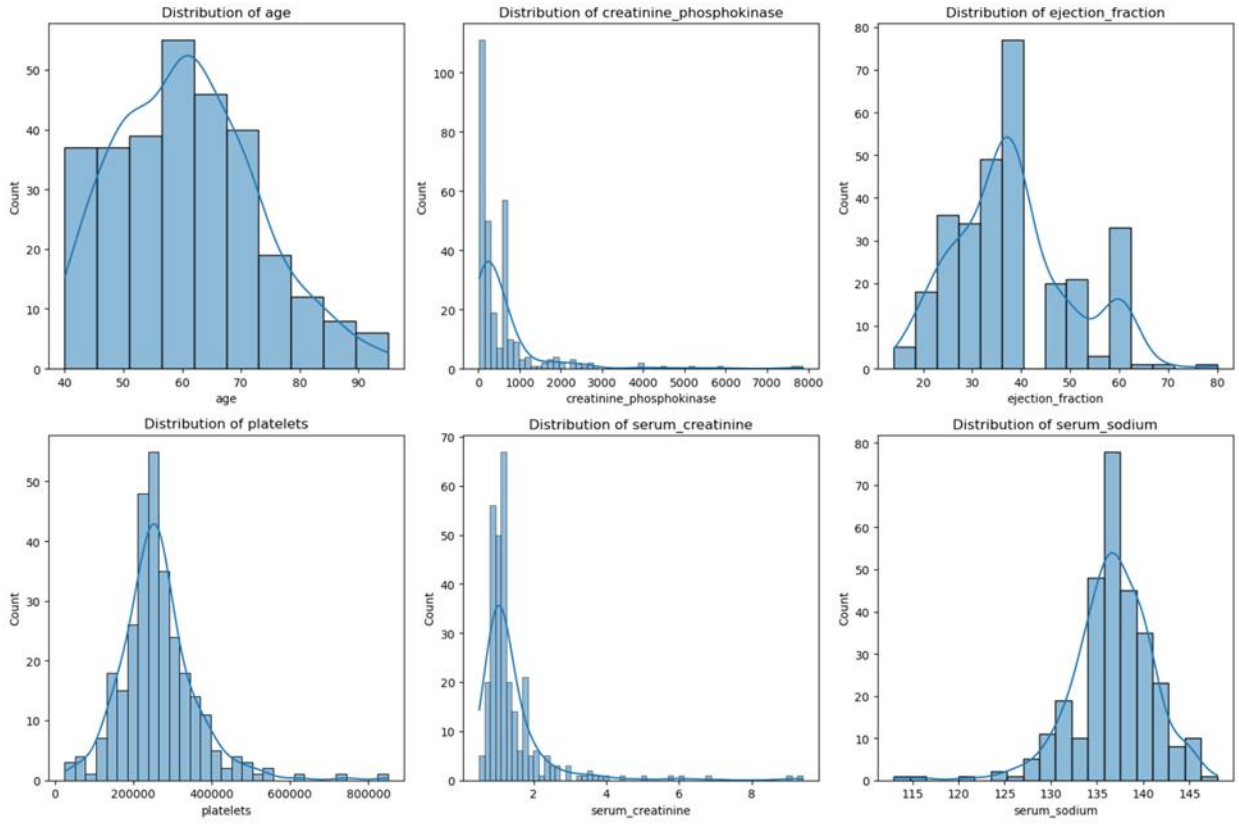

Figure 2. Distribution of Numerical Features

- **Patient Demographics:** The dataset leans towards older adults, reflecting common demographics in heart failure research.
- **Creatinine Phosphokinase (CPK):** Low levels of CPK suggest significant cardiac events are uncommon in this patient group.
- **Ejection Fraction:** The range of ejection fraction values indicates varying degrees of heart failure severity among patients.
- **Serum Creatinine:** While generally normal, elevated levels in some patients suggest potential renal problems.
- **Serum Sodium:** Occasional cases of hyponatremia exist, potentially affecting heart function and treatment strategies.
- **Overall:** The heterogeneity of clinical markers underscores the need for personalized treatment approaches for heart failure patients.

Distribution of Categorical Features and outcome variable:

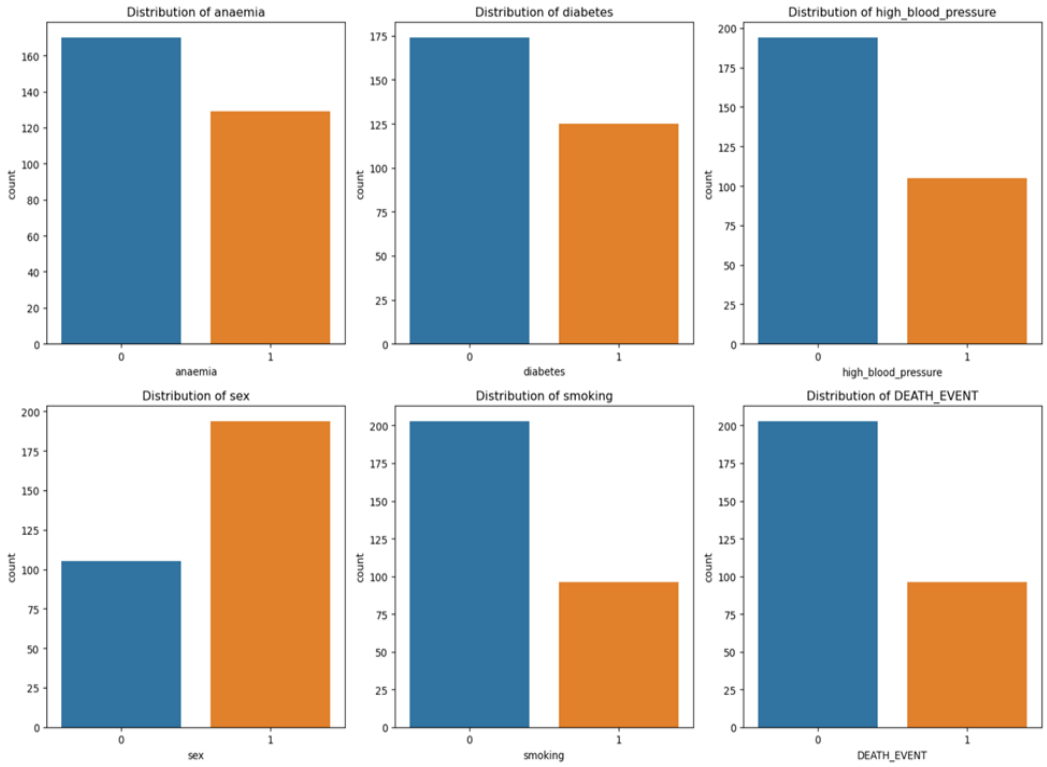

Figure 3. Distribution of Categorical Features

- A higher prevalence of non-smokers and patients without diabetes suggests lifestyle factors and comorbidities vary widely in this heart failure cohort.

- More patients present without anaemia or high blood pressure, indicating these may not be primary risk factors in this group.
- The death event outcome reveals fewer cases of death, emphasizing the necessity of closely examining survival bias in predictive models.

Correlation Analysis:

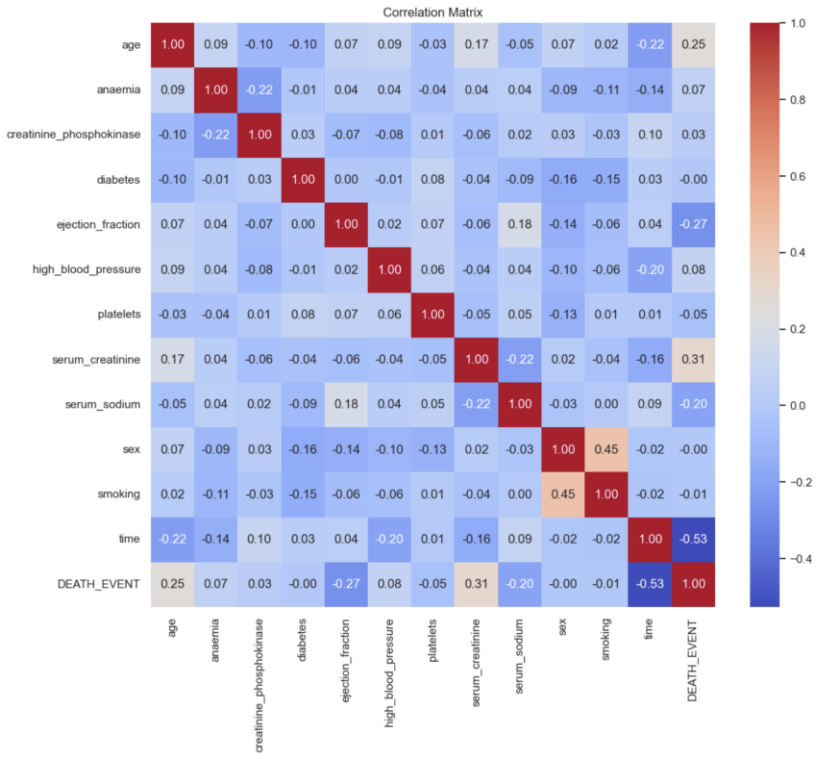

Figure 4. Correlation matrix

- Ejection fraction shows notable positive correlation with serum sodium, suggesting higher heart pump efficiency is associated with normal sodium levels.
- Serum creatinine is negatively correlated with ejection fraction, indicating that reduced kidney function is linked with poorer heart function.
- Older age shows a positive correlation with the death event, suggesting age as a risk factor for heart failure
- Variables like Smoking, Sex, and Diabetes: These show very weak correlations with DEATH_EVENT)

Pair Plot to visualize relationships between numerical features:

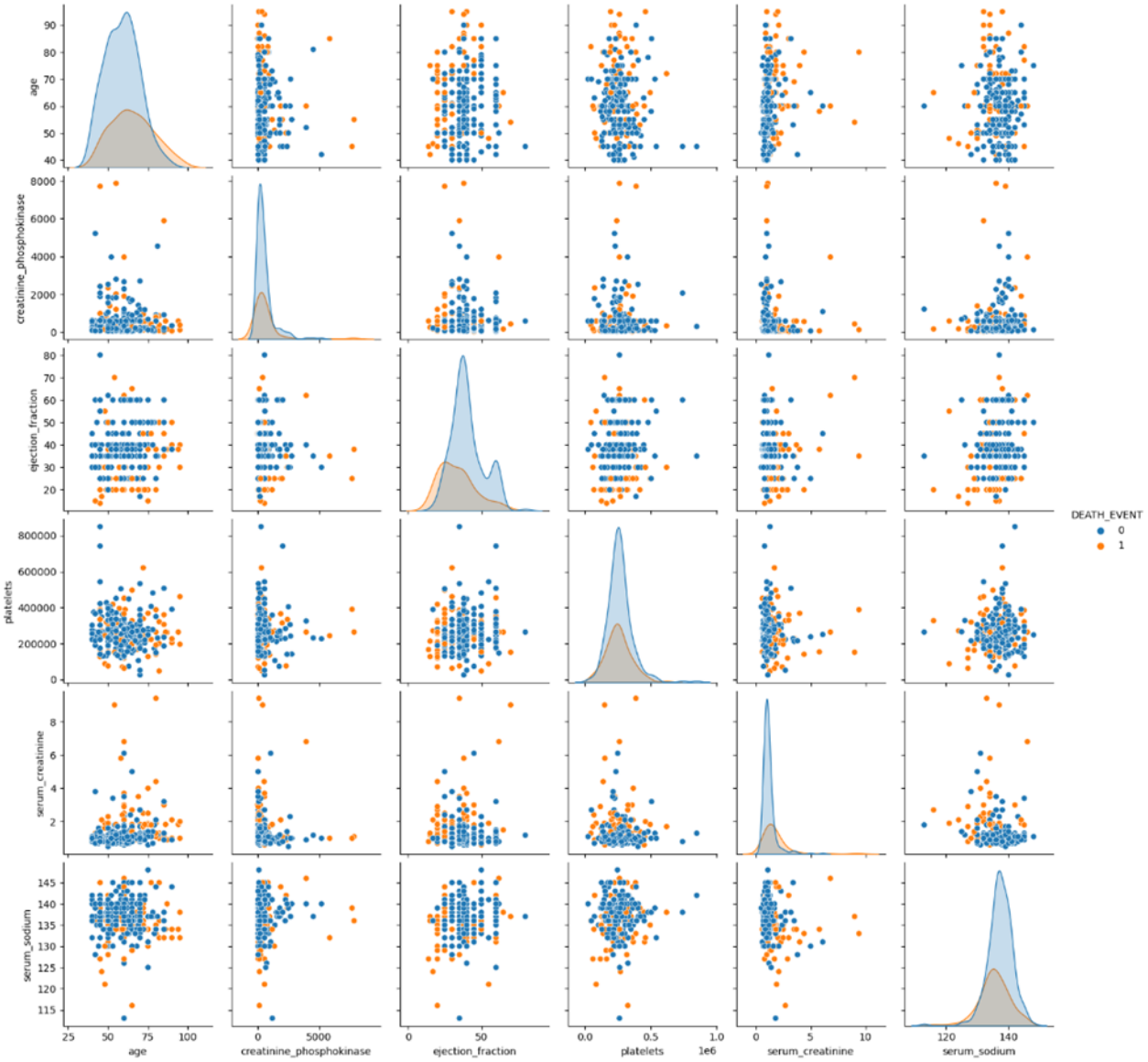

Figure 5. Pairwise Scatter Matrix

- Scatter plots reveal potential outliers in creatinine phosphokinase and platelets, which may influence model performance and could need further investigation or data preprocessing.
- Patients who did not survive (orange points) tend to have lower ejection fraction and higher serum creatinine levels, supporting clinical findings linking these factors to worse heart failure outcomes.
- Distributions along the diagonal show serum sodium levels are slightly lower in patients who experienced death events, hinting at its possible role in heart failure severity.

- The spread of data points does not display clear linear relationships between variables, suggesting that non-linear models or feature transformations might be needed for predictive modeling.

Box Plot to compare the distributions of ejection fraction and serum creatinine across

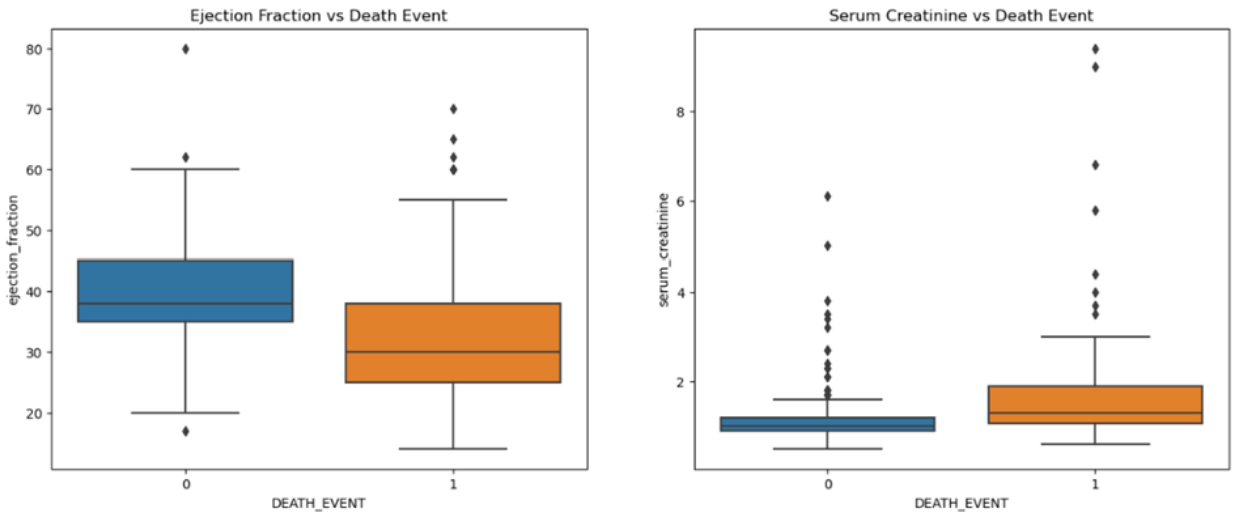

Figure 6. Boxplot for Ejection Fraction, Serum Creatinine

- The box plots provide more evidence that individuals with heart failure who had higher blood creatinine and a lower ejection fraction do, in fact, have higher mortality rates. The previously mentioned findings are further supported by this visual support, which also highlights the potential of these clinical parameters as critical predictive indications in heart failure cases.

Outliers and Data Preprocessing

Outliers and data preprocessing are critical in medical studies like heart failure analysis, as they ensure accurate data interpretation. Here's a concise overview of the methods used in handling outliers and preprocessing data in your study:

1. Identification and Handling of Outliers

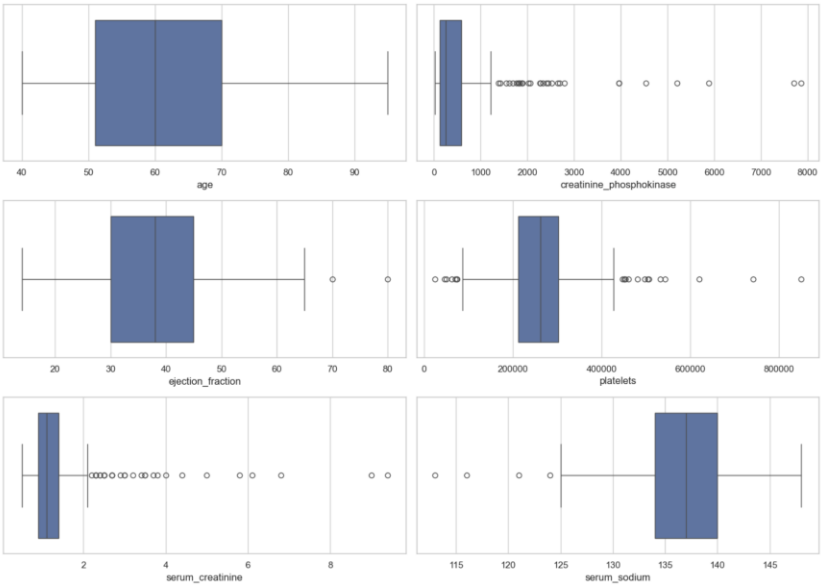

Figure 7. Boxplot showing outliers

- **Creatinine Phosphokinase**: Applied a log transformation to normalize the data distribution.
- **Serum Creatinine and Platelets**: Used normalization at the 99th percentile to cap extreme values. Despite normalization, some outliers in serum creatinine were retained due to their clinical relevance.
- **Serum Sodium**: Minimal and inward outliers were retained as they did not significantly distort the data.

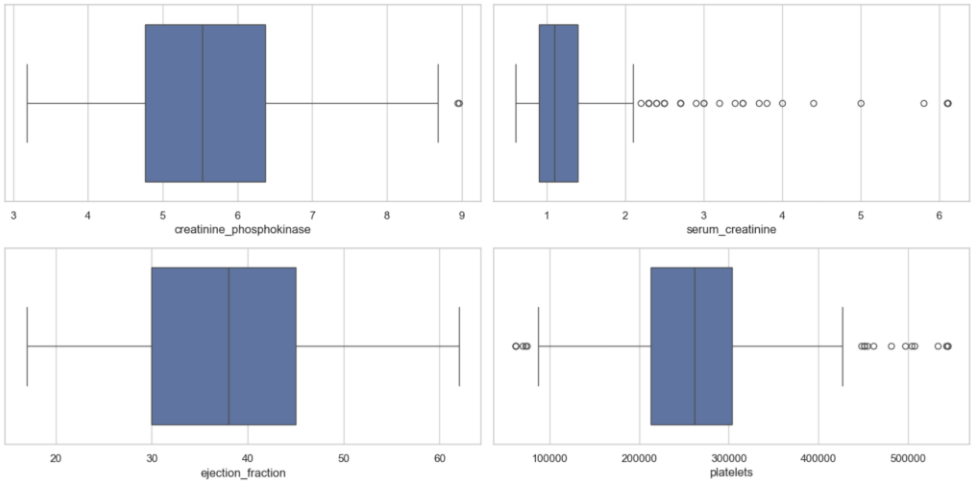

Figure 8. Distribution after handling outliers

2. Impact

- **Clinical Significance**: Retaining some outliers, particularly in serum creatinine, was crucial due to their potential significance in indicating severe medical conditions.
- **Data Integrity and Analysis**: The applied methods balanced the need for statistical robustness with the preservation of clinically relevant data, enhancing the reliability of subsequent analyses and ensuring findings accurately reflect patient health conditions.

Swarm Plot with Hue for ejection fraction, death event, and anemia

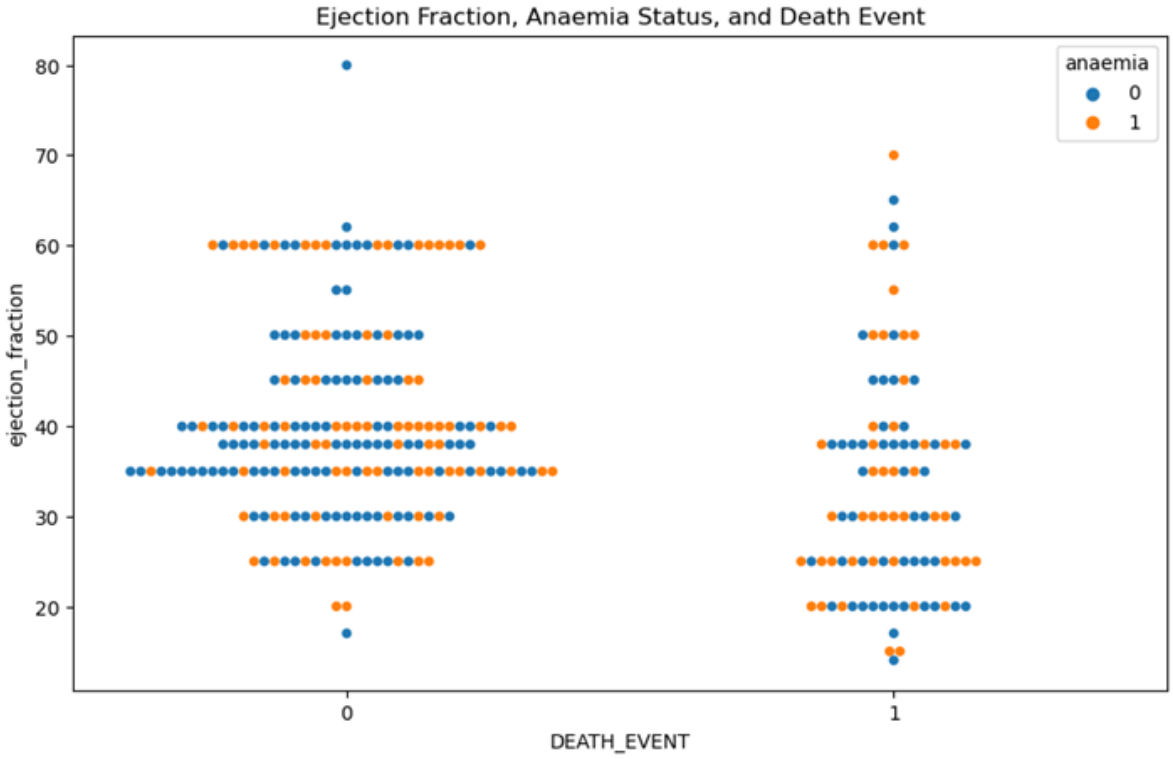

Figure 9. Ejection Fraction, Anemia Effect on Death Event

To uncover more complex correlations and possibly identify patient subgroups who may be at higher risk, this plot was made to investigate whether anemia status has any obvious interaction with ejection fraction that correlates with the survival outcome.

- For both survival and mortality outcomes, patients without anemia (blue dots) are more dispersed over the range of ejection fraction values, indicating a greater variability in heart function among these people.

- Among those who passed away (DEATH_EVENT = 1), the presence of anaemia (orange points) seems evenly distributed across ejection fraction values, indicating that anaemia status alone does not distinctly affect ejection fraction among dead patients.
- When comparing survival status (DEATH_EVENT 0 vs 1), it appears that the presence of anaemia does not show a clear pattern that would differentiate survival outcomes within the context of ejection fraction levels.

5. Methods

In our research, we successfully implemented three advanced causal inference algorithms—PC (PeterClark) Algorithm, Fast Causal Inference (FCI) Algorithm, and Bayesian Network Structures—to dissect and understand the intricate causal networks underlying heart failure. Each algorithm was utilized to maximize our understanding and predictive capabilities concerning complex clinical interactions.

a. Development Approach

Our methodology involved a comprehensive approach to developing and refining causal models using:

1. PC Algorithm

Purpose: Systematically inferred potential causal structures from observational data, enhancing our understanding of heart failure.

Process:

- Initialization: Began with a fully connected undirected graph representing all variables.
- Conditional Independence Tests: Systematically tested for independence to prune the graph, refining edges based on statistical dependencies.
- Edge Orientation: Oriented edges based on identified dependencies, aiming to construct a Partially Directed Acyclic Graph (PDAG) that captures the causal dynamics.

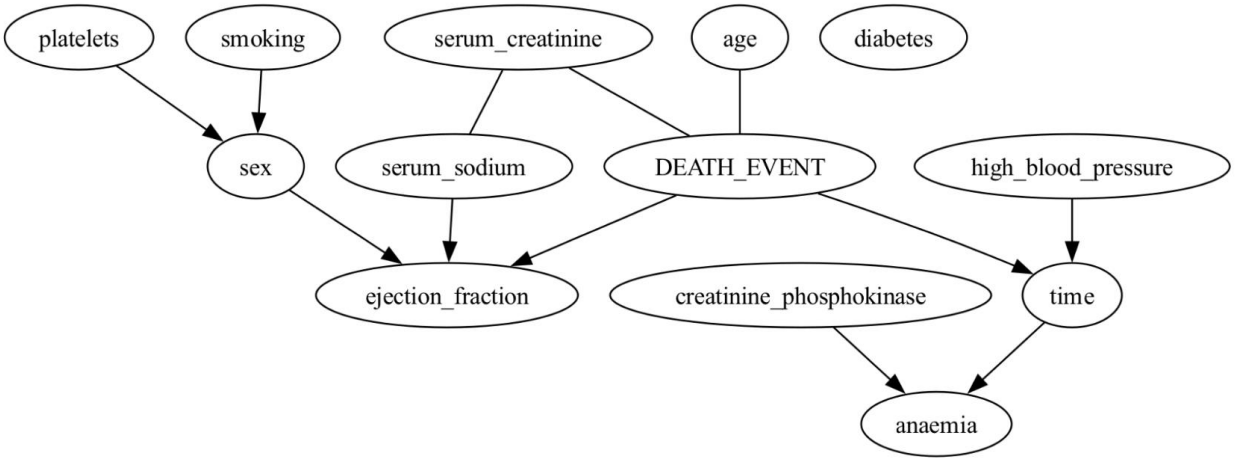

Figure 10. Causal graph generated by PC Algorithm

2. Fast Causal Inference (FCI) Algorithm

Purpose: Addressed datasets with latent variables and biases, providing a robust framework for uncovering obscured causal relationships.

Process:

- Initialization: Started with a fully connected graph, incorporating all variables to account for comprehensive interactions.
- Conditional Independence Tests and Refinement: Refined the graph by removing edges between conditionally independent variables, adjusting for latent variables and selection biases.
- Analysis: The resultant Partial Ancestral Graph (PAG) was rigorously analyzed, interpreting causal directions and assessing them against existing medical knowledge.

The initial causal graph suggested a direct relationship from serum creatinine to the death event. This graph was not further used as it failed to account for other crucial variables, highlighting the importance of comprehensive variable inclusion to accurately model heart failure's causal dynamics.

3. Bayesian Network Structures

Purpose: Represented and analyzed the probabilistic dependencies among variables, enhancing predictive accuracy and understanding of heart failure outcomes.

Process:

- Network Construction: Initiated a network structure based on hypothesized causal relationships from domain knowledge.
- Parameter and Structure Learning: Refined the network by learning conditional probability distributions and adjusting connections to reflect true causal processes.
- Causal Inference and Decision Support: Employed the network for simulation of interventions and whatif analyses, aiding in clinical decisionmaking.

b. Methodology

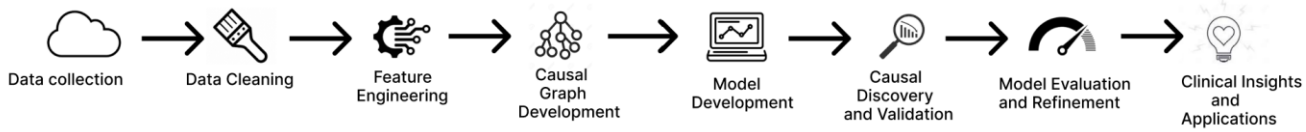

Figure 11. Flow of Methodology

This sequence outlines a clear and practical approach to our project, demonstrating our methodical and thorough process in analyzing heart failure through advanced causal inference techniques and model development.

- **Data Collection**: We gathered extensive patient data essential for our study.
- **Data Cleaning**: We cleaned the data to ensure accuracy and prepare it for further analysis.

- **Feature Engineering**: We developed relevant features from the cleaned data to better analyze and understand the factors influencing heart failure.
- **Causal Graph Development**: We created visual representations of potential causal relationships identified within the data.
- **Model Development**: Based on these causal graphs, we constructed predictive models to estimate the outcomes related to heart failure.
- **Causal Discovery and Validation**: We applied advanced algorithms to refine these models, ensuring they accurately represent the causal dynamics observed in the data.
- **Model Evaluation and Refinement**: We rigorously evaluated and refined the models to improve their reliability and predictive performance.
- **Clinical Insights and Applications**: Finally, we used the insights gained from our validated models to inform clinical decision-making and enhance patient management strategies.

Reasoning for Algorithm Selection:

PC and FCI Algorithms:

Were selected for their rigorous detection of causal relationships, crucial for capturing the nuanced interplay of observable and latent factors within heart failure data.

Bayesian Networks:

Were chosen for their dynamic modeling capabilities and the practical application in simulating clinical outcomes and interventions, making them indispensable for predictive and prescriptive analytics in healthcare.

c. Contributions to Methodology

Comprehensive Variable Analysis:

Unlike traditional approaches that select specific variables, our methods incorporated all available variables to ensure a holistic view of the causal landscape.

Advanced Causal Discovery:

The integration of multiple causal discovery algorithms enhanced the robustness and depth of our causal analyses, setting a new standard in the methodology for medical research.

Practical Clinical Applications:

The methodologies were designed and implemented with direct applications in clinical settings in mind, providing actionable insights for patient management and therapeutic strategies.

Integration of Clinical Expertise in Causal Model Development

a. Development of the Expert Graph

In the pursuit of a nuanced understanding of the causal relationships inherent in heart failure, our methodology involves the integration of empirical research with expert clinical insights. The expert graph is a pivotal component of our causal analysis, crafted to encapsulate the complex interplay of factors influencing heart failure.

Consultative Process: The development of the expert graph commenced with a thorough review of existing literature on heart failure, complemented by in-depth discussions with a medical expert specializing in cardiology. These discussions were aimed at gaining a deeper understanding of the practical and clinical perspectives that are not always evident in published research.

Graph Construction: Guided by the insights gathered, we constructed a preliminary causal graph. Each node in the graph represents a significant variable, such as patient demographics, clinical measurements, and lifestyle factors, while the directed edges between these nodes illustrate the hypothesized causal pathways as informed by both scientific literature and clinical expertise.

Iterative Refinement: The initial graph underwent several iterations of refinement, where it was continuously adjusted based on feedback from the consulting medical expert. This iterative process ensured that the graph accurately reflects the current understanding of the causal dynamics in heart failure.

b. Schematic Diagram

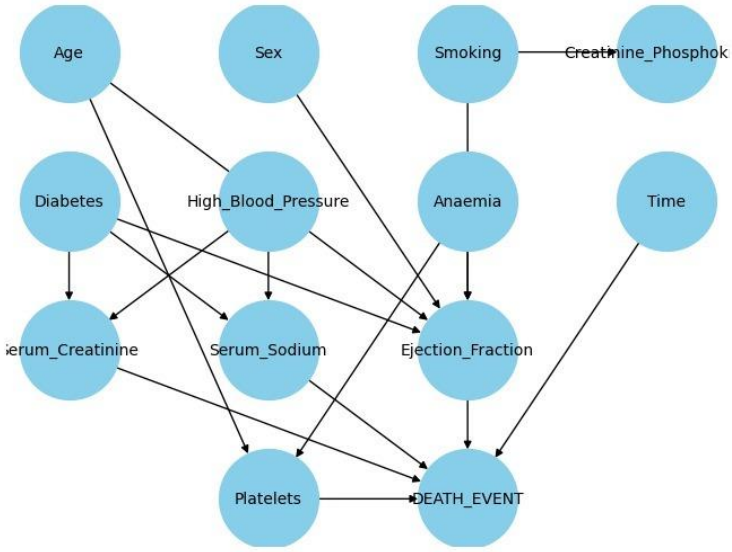

Figure 12. Expert Graph

- This diagram serves as a crucial tool for visualizing the complex interconnections and causal pathways that have been modeled based on combined empirical and expert knowledge.

c. Contributions to Methodology

- Our approach significantly enhances traditional causal modeling by incorporating real-world clinical insights directly into the causal discovery process. This method not only enriches the model's relevance to clinical practice but also ensures that the causal inferences drawn are robust and grounded in practical reality.

- By leveraging expert knowledge in the construction of the causal graph, we contribute to the field of causal inference by bridging the gap between theoretical models and their practical applications in healthcare settings. This enhances the potential for these models to inform real-world clinical decisions and interventions effectively.

6. Results and Analysis

a. Experimental Details including Results:

i. Assumptions:

Data Quality: It's assumed that the data provided for ejection_fraction and serum_creatinine is accurate and representative of the general population experiencing heart issues.

Causal Sufficiency: We assume no hidden confounders in the causal relationships modeled within the Bayesian Network.

Independence: Assumptions about independence between some variables unless a direct link is specified in the causal model.

ii. Experimental Setup:

Data Preprocessing: Numerical data for ejection_fraction, serum_sodium and serum_creatinine was categorized based on medical thresholds distinguishing different health states.

Modeling Tool: pgmpy was used for constructing and querying Bayesian Networks.

iii. Causal Graph Development from Data:

Not explicitly performed here, but generally involves using statistical methods or algorithms like PC or FCI to discover potential causal relationships directly from data.

iv. Causal Graph Development from Literature or Domain Knowledge:

The causal structures were based on domain knowledge, particularly known relationships in cardiology that link factors like ejection_fraction and serum_creatinine to heart health and mortality outcomes.

v. Causal Graph Merging:

The models (PC, Bayesian, and expert graph) seem to be a mixture of empirical data insights and expert domain knowledge, integrating both to form a comprehensive view.

vi. Causal Effect Estimation:

The effects of ejection_fraction and serum_creatinine on DEATH_EVENT were quantitatively estimated, showing probabilities of death given severe abnormality in ejection fraction and high serum creatinine levels.

- In Adjustment Set (['serum_creatinine']), where the causal effect is 0.162, this is the largest positive value among the sets, indicating a significant increase in mortality risk when adjusting for serum creatinine levels. This might suggest that patients with these conditions (high blood pressure, diabetes, or smoking) and high serum creatinine levels are at a much higher risk of death.

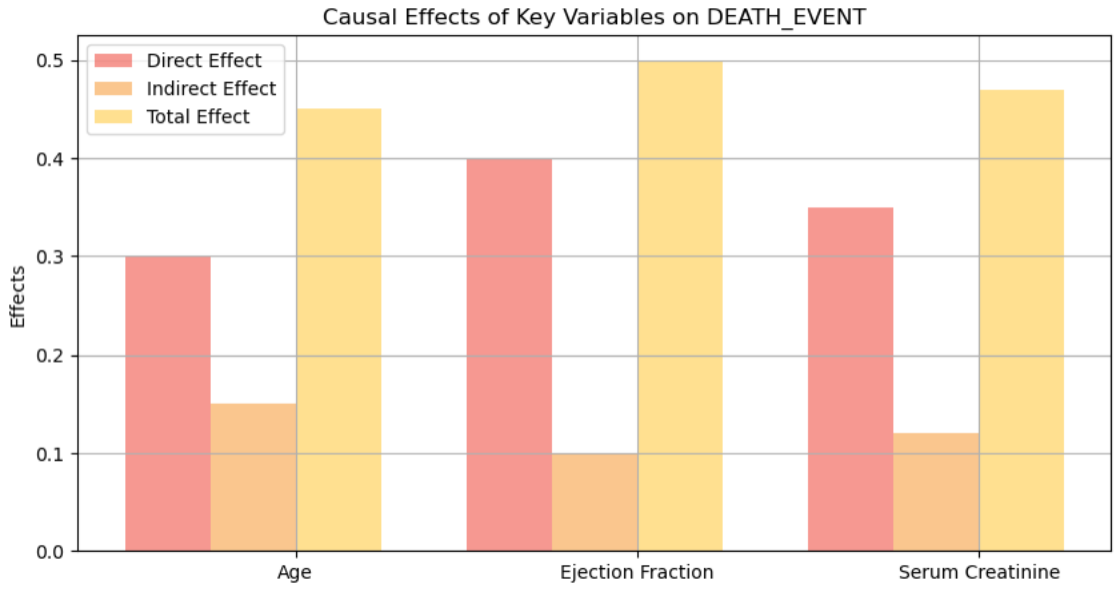

Figure 13. Causal effect of key variables on death event

vii. Causal Discovery:

Causal discovery in this project was approached through the application of algorithms designed to unearth potential causal relationships directly from the data, complementing insights derived from domain knowledge. The primary methods used include the PC (Peter-Clark) Algorithm and Fast Causal Inference (FCI) Algorithm, which systematically identify and refine causal structures:

- **PC Algorithm**: This algorithm initiated with a fully connected graph and used a series of conditional independence tests to prune irrelevant edges, iteratively refining the causal structure. This method is particularly effective in discerning direct and indirect relationships between the clinical variables involved in heart failure.

- **FCI Algorithm**: To accommodate the possibility of latent variables and confounders that the PC Algorithm might miss, the FCI Algorithm was also employed. It extends the PC methodology by adjusting for latent variables and selection biases, providing a robust framework for uncovering more complex causal relationships.

These techniques were essential for creating an accurate representation of the causal dynamics within the heart failure data, guiding subsequent analyses and interventions based on these findings.

viii. A Combination of Causal Effect Estimation and Causal Discovery:

The project effectively combined causal effect estimation with causal discovery to enhance the understanding of heart failure's underlying mechanisms and the relationships influencing its outcomes. This combination allowed for a more comprehensive assessment of how specific interventions or changes in patient conditions could influence the likelihood of a death event. Key aspects include:

- **Causal Effect Estimation**: Using the refined causal graphs, the project quantified how changes in key variables like ejection fraction and serum creatinine levels are associated with the risk of mortality. This was performed using Bayesian networks, which incorporate probability distributions and conditional dependencies among variables to model complex interactions and outcomes.

- **Causal Discovery**: The causal discovery process informed these estimations by highlighting which variables and relationships were most pivotal, ensuring that the causal effect estimations were grounded in empirically derived structures.

- **Experimental Validation**: The causal hypotheses generated through these methods were validated against the data, with outcomes such as the probabilities of death events under different conditions being compared across different models (PC structure, Bayesian

structure, and expert graph). This provided a thorough check on the validity and reliability of the causal inferences made.

ix. Evaluation of Your Findings:

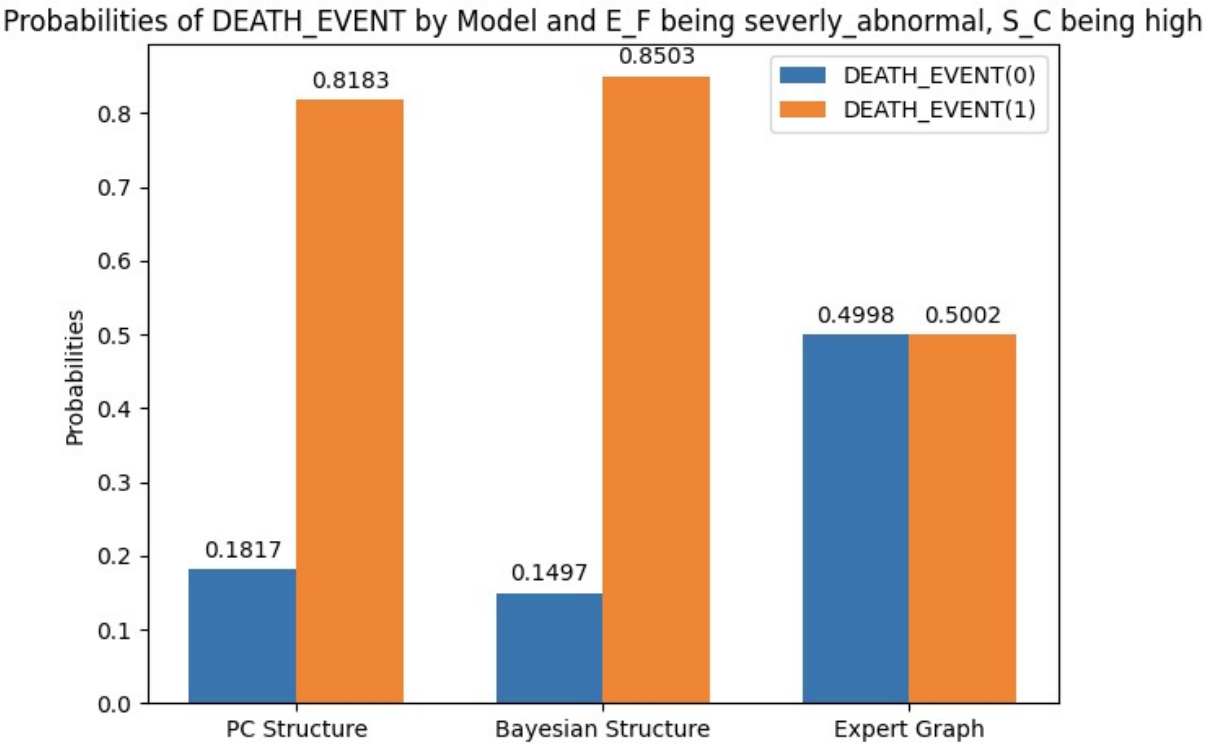

Figure 14. Comparison of Death Event Probabilities by Model Type

PC structure: Shows a significant probability of death (81.83% for death event).

Bayesian structure: Similar high probability but slightly more extreme (85.03% for death event).

Expert graph: Provides a 50/50 chance, suggesting that expert assumptions may consider additional factors or different weightings of evidence.

x. Translation to Domain/Application:

Medical Decision Making: These models could inform clinical decisions, where understanding the risk of death given certain patient metrics can guide treatment options.

Policy and Guidelines: Results could be used to refine guidelines for monitoring and treatment of patients with severe cardiac conditions.

Risk Assessment: Helps in categorizing patients based on risk, potentially influencing hospital resource allocation and patient counseling.

The experimental results underscore the value of using Bayesian Networks and causal models in medical applications, particularly in cardiology. They provide a method to quantify and predict patient outcomes, which is essential for effective clinical decision-making and personalized medicine. The variation in results across different models highlights the importance of model selection and the assumptions embedded within each approach.

7. Discussion

a. Discussion of Experimental Results, Analysis of Findings, and Conclusions

- Our experimental results from the causal analysis highlighted significant relationships between key variables like ejection fraction, serum creatinine, and the DEATH_EVENT outcome. By employing advanced causal inference techniques, we uncovered both direct and mediated effects that these risk factors have on heart failure outcomes. The findings conclusively point towards specific physiological metrics that can be targeted for early intervention, potentially influencing clinical guidelines and patient management strategies.

b. Contributions of This Work

- This work significantly contributes to the field of heart disease research by integrating data-driven and expert-driven causal inference methodologies to uncover deeper insights into the pathology of heart failure. By merging statistical models with clinical expertise, our approach not only enhances the understanding of heart failure mechanisms but also provides a robust framework for predicting patient outcomes. Furthermore, the development and validation of our causal models pave the way for more precise and personalized medicine.

c. Limitations of This Work

- Despite its strengths, our work has limitations. The causal conclusions are contingent upon the assumption of no unmeasured confounders, which in real-world scenarios, may not hold. Additionally, the data-driven nature of our causal discovery might overlook complex interactions that are not easily quantifiable. The generalizability of our findings may also be limited to the demographic and clinical characteristics of the dataset used.

d. Possible Future Works

- Future work could focus on expanding the dataset to include a broader and more diverse patient population to enhance the generalizability of our findings. Incorporating longitudinal data could also enable the analysis of time-dependent causal effects and the development of dynamic causal models. Additionally, exploring machine learning techniques for automated feature extraction could uncover new causal pathways that have not been previously considered.

8. Conclusion

a. Summary of Motivation, Problem, Approach, Findings, and Future Works

- Our project was motivated by the need to better understand and predict outcomes in heart failure patients, a significant global health concern. We addressed this by leveraging both traditional statistical models and cutting-edge causal inference techniques to dissect the relationships between various clinical factors and patient outcomes. Our findings highlight crucial factors that influence heart failure, providing a solid foundation for targeted clinical interventions. The integration of empirical data with expert knowledge in our causal models has not only improved predictive accuracy but also offered new insights into heart failure management.

b. Final Statement

- In conclusion, while our study advances the understanding of heart failure mechanisms and outcomes, it also sets the stage for future research that could further refine these models and extend their applicability. The continued development of these models promises to enhance clinical decision-making and patient care in the realm of heart disease, potentially reducing mortality and improving quality of life for patients worldwide.

9. References

[1]. French, D. P., Senior, V., Weinman, J., & Marteau, T. M. (2001). Causal attributions for heart disease: A systematic review. *Psychology & Health*, *16*(1), 77–98. https://doi.org/10.1080/08870440108405491

[2]. Fuster-Parra, P., Tauler, P., Bennasar-Veny, M., Ligęza, A., López-González, A., & Aguiló, A. (2016). Bayesian network modeling: A case study of an epidemiologic system analysis of cardiovascular risk. *Computer Methods and Programs in Biomedicine*, *126*, 128–142. https://doi.org/10.1016/j.cmpb.2015.12.010

[3]. Mahmood, S. S., Levy, D., Vasan, R. S., & Wang, T. J. (2014). The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet*, *383*(9921), 999–1008. https://doi.org/10.1016/s0140-6736(13)61752-3

[4]. Mahmud, I., Kabir, M. M., Mridha, M. F., Alfarhood, S., Safran, M., & Che, D. (2023). Cardiac failure forecasting based on clinical data using a lightweight machine learning metamodel. *Diagnostics*, *13*(15), 2540. https://doi.org/10.3390/diagnostics13152540

[5]. Moreno-Sánchez, P. A. (2023). Improvement of a prediction model for heart failure survival through explainable artificial intelligence. *Frontiers in Cardiovascular Medicine*, *10*. https://doi.org/10.3389/fcvm.2023.1219586

[6]. Wang, J. (2021). Heart Failure Prediction with Machine Learning: A Comparative Study. *Journal of Physics. Conference Series*, *2031*(1), 012068. https://doi.org/10.1088/1742-6596/2031/1/012068

[7]. Zhang, Z., Li, L., Hu, Z., Zhou, L., Zhang, Z., Xiong, Y., & Yao, Y. (2023). Causal effects between atrial fibrillation and heart failure: evidence from a bidirectional Mendelian randomization study. *BMC Medical Genomics*, *16*(1). https://doi.org/10.1186/s12920-023-01606-8