UNIVERSITY
OF HULL

Progression of risk in heart failure using dynamic risk modelling

being a Thesis submitted in partial fulfilment of the requirements

for the Degree of Doctor of Philosophy in Department of Computer Sciences

at the University of Hull

by

Syed Mohsin S. Kazmi
BSc, MCS & IT, MSc Network Centred Computing

Principal Supervisor:
Dr. Chandrasekhar Kambhampati

Clinical Supervisor
Prof. Andrew L. Clark

May 2024

To my beloved family, especially to my brother who was taken from us too soon by heart failure

Acknowledgements

My journey through my PhD has been a whirlwind, particularly over the last few years, which have been a critical and extremely challenging. During this time, I experienced the loss several beloved individuals, making it incredibly difficult to concentrate and maintain focus on my studies. Their absence deeply affected me adding an emotional strain to an already challenging academic journey.

My background in IT and diverse clinical research experience has empowered me to apply novel yet intricate concepts into practical solutions. I am forever grateful for the opportunity presented to me, and I am determined to continue on this path with success. When I first began working on my PhD proposal, I initially grappled with uncertainty regarding how to achieve my aims and objectives I set for research. However, collaboration with talented professionals from various fields has enriched my perspective and helped me stay in the right direction. Their support and insights were invaluable, helping me stay on track even with I doubted myself. Embracing this challenge brings me immense joy and fulfilment.

With deep appreciation, I extend my heartfelt gratitude to Dr. Chandrasekhar Kambhampati whose patience, encouragement, and unwavering guidance have been instrumental throughout my journey. Without his invaluable support, completing this thesis would not have been possible. I am profoundly grateful to Professor Andrew Clark for his immense support during my PhD. His expertise provided me with confidence and tools needed to negate the complexities of my research. Mentorship I received from Dr. Kambhampati and Prof Clark has not only shaped my academic growth but also profoundly influenced my personal development. They helped me in every stage of PhD, and I am particularly grateful for their support when I was at my lowest around the time of the loss of my loved ones. I also

want to express my sincere thanks to Prof. John GF Cleland for his support whenever I sought his help. His willingness to assist and his valuable insights have been crucial in helping me overcome various challenges during my research. I am particularly thankful to be constantly surrounded by amazing people. I'd like to thank Prof Alyn Morice, Wayne Sheedy, Rachel Thompson and other members of Cardiology and Respiratory research teams for helping me along this journey. I am grateful to Professor Morice for allowing me the time to finish my PhD.

I dedicate my PhD to my loving parents, Abbu Ji (Syed Abdul Hussain Shah Kazmi and Ammi Ji (Syeda Zakia Kazmi) and brother Bhaiya (Syed Zahid Hussain Kazmi). Though they are no longer with me physically, their unwavering belief in me and my aspirations throughout my life has been a source of great strength. If they were alive today, I know they would be incredibly proud and overjoyed by this achievement. I take comfort in the thought that they are watching over me from the heavens, sharing in my happiness and success.

I am also grateful to my dear sisters, Syeda Tasneem Kazmi, Syeda Fozia Kazmi, and my brothers, Eng. Syed Shahid Hussain Kazmi and Dr. Syed Khurran Kazmi. Their continuous support and encouragement have been invaluable, without which I could not have reached this significant milestone. Their presence and reassurance during challenging times reminded me of the importance of balance beyond academia. Words cannot fully express how much their support has meant to me, and I am deeply thankful to have such wonderful siblings by my side.

I want to express my heartfelt gratitude to my beloved wife, Syeda Saleha. From the last few years journey, when everything was so tough for me you've been there for me every step of the way. Your unwavering support has been my greatest strength during these years. You've brought me immense happiness and peace, even during the most stressful times.

With you all by my side, I feel like I can achieve anything. Thank you for everything, my dears.

Publications and Conferences

I hereby confirm that the material in this thesis is the result of my original work. Portions of this research have been presented at various conferences and published in academic journals. During the period of my PhD, I have more than 40 publications including abstracts. The list of publication can be seen on https://scholar.google.co.uk/citations?user=5lIT1m8AAAAJ&hl=en

The following are directly relevant to my PhD

Publications

Kazmi S, Kambhampati C, Cleland JGF, Cuthbert J, Kazmi KS, Pellicori P, Rigby AS, Clark AL. Dynamic risk stratification using Markov chain modelling in patients with chronic heart failure. ESC Heart Fail. 2022 Oct;9(5):3009-3018. doi: 10.1002/ehf2.14028. Epub 2022 Jun 23. PMID: 35736536; PMCID: PMC9715820.

Kazmi S, Kambhampati C, Cleland JGF, Cuthbert J, Kazmi KS, Pellicori P, Rigby AS, Clark AL. Disease progression in chronic heart failure is linear. Insights from multistate modelling. EURJHF. European Journal of Heart Failure.

Conference Abstracts

Absorbing Markov Chains for modelling progress of patients with heart failure: a case study of Hull Life Lab Authors: **S Kazmi** (Hull,GB), J G Cleland (Glasgow,GB), C Kambhampati (Hull,GB), P Pellicori (Glasgow,GB), A Rigby (Hull,GB), A L Clark (Hull,GB) ESC - European Society of Cardiology, 2020

Tracking disease progression in patients with heart failure using a Markov chain
Model Authors: **S Kazmi** (Hull,GB), J G F Cleland (Glasgow,GB), C Kambhampati (Hull,GB), P Pellicori (Glasgow,GB), A S Rigby (Hull,GB), A L Clark (Hull,GB) ESC - European Society of Cardiology, 2021

Dynamic risk prediction in heart failure using absorbing Markov chains
Model Authors: **S Kazmi** (Hull,GB), J G F Cleland (Glasgow,GB), C Kambhampati (Hull,GB), P Pellicori (Glasgow,GB), A S Rigby (Hull,GB), A L Clark (Hull,GB) BSH British society of Heart Failure, 2020

List of Abbreviations

ACE	Angiotensin-converting enzyme inhibitor
6 Min	6 Minutes' walk test
AI	Artificial Intellegence
AMC	Absorbing Markov chains
ARB	Angiotensin receptor blocker
Bb	Beta-blocker
BL	Baseline
BMI	Body mass index
BP	Blood pressure
bpm	Beats per minute
BSE	British Society for Echocardiography
CHF	Chronic heart failure
CUR	Current
CV	Cardiovascular
Demo	Demography
E.g. ,	For example
ECG	Electrocardiogram
Echo	Echocardiogram
EF	Ejection fraction
eGFR	Estimated glomerular filtration rate
ESC	European Society for Cardiology
FU	Follow up
HeFNEF	HF with a normal (or "preserved") ejection fraction
HeFPEF	HF with preserved left ventricular ejection fraction
HeFREF	HF with reserved ejection fraction
HES	Hospital episode statistics
HEY	Hull East Yorkshire
HF	Heart failure
HLL	Hull LifeLab

HR	Hazard ratio
i.e.,	For that
IQR	Interquartile range
kNN	k-nearest neighbour algorithm
LOCF	Last Observation Carried Forward
LVI	Left ventricular impairment
LVSD	Left ventricular systolic dysfunction
MAR	Missing at random
MC	Markov chain
MCAR	Missing completely at random
ML	Machine learning
MNAR	Missing not at random
MRA	Mineralocorticoid antagonist
MRI	Magnetic resonance imaging
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NT-proBNP	N-terminal pro–B-type natriuretic peptide
NYHA	New York Heart Association class
ONS	Office for National Statistic
OPD	Out-patient department
PDM	Patient data models
PH	Cox proportional hazards (PH) regression
PPS	Palliative performance scales
QoL	Quality of Life
RMC	Regular Markov chains
SBP	Systolic blood pressure (SBP)
SD	Standard deviation
SVM	Support Vector Machine
TI	Transition intensities
TP	Database table

Abstract

Heart failure (HF) is a prevalent condition affecting a significant number of individuals in the UK, leading to substantial healthcare utilisation and adverse outcomes. Despite advancements in treatment and management, the prognosis for hospitalised HF patients remains poor, with a one-year mortality rate of 40%. Improving risk modelling and predictive assessment is crucial for enhancing patient outcomes and reducing healthcare burden.

This thesis aims to analyse HF progression, improve data handling for risk analysis, and develop machine learning (ML) models for tracking health status changes and predicting risks over the course of the disease. To do this a dynamic risk modelling approach was developed and used. This started with the use of a Naïve modelling process with standard ML methodology to the use of Markov chains (MC) and multistate modelling (MCM) with modified MC. The models incorporated dual temporal perspective to investigate the progression of risk which comprises of both short-term and long-term prediction, enabling more accurate forecasting of patient outcomes across varying timeframes.

It was found that the MSM methods can predict a) hospitalisation and mortality both at population level and individual level b) can also determine the number of time visit are made to both [*Hosp*] and [*OPD*] before patient died. An expected finding of this thesis is that the progression of HF is linear and not non-linear as it has been assumed. At the same, the modelling in this thesis contributed valuable insights into the progression of risk in HF and underscored the importance of dynamic risk modelling for prognostic assessment. Recommendations for future research include further validation and refinement of the model to enhance predictive accuracy and clinical utility.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation and research problems

Chronic heart failure (CHF) is widespread and consumes a lot of healthcare resource (Colombo et al., 2008; Ketchum & Levy, 2011; Hutchinson et al., 2014; National Heart Failure Audit Report UK, 2019; Taylor et al., 2019; Abel et al., 2024). Patients with CHF have a high mortality and are admitted to hospital frequently (Roger, 2013; Shiraishi et al., 2018). The most significant contributor to the cost of treatment for CHF is hospitalisation (Shoaib et al., 2016; Cuthbert et al., 2024). However, it can be challenging to interpret the epidemiological data on heart failure (HF) and its exact scale since there is lack of a definitive gold standard for diagnosing HF (Futoma et al., 2015). Currently, most modelling efforts are focused on applying scoring systems to assess the risk of death for individual patients, which might be helpful for that patient but it does not capture the patterns of disease behaviour at a population level (Ieva et al., 2017; Levin et al., 2018).

HF is a complex clinical syndrome that results from any structural or functional impairment of ventricular filling or ejection of blood (NICE, 2010; Cleland et al., 2012-13). Patients with HF can experience breathlessness and fatigue on exertion, together with ankle swelling (with generalised swelling due to fluid retention being the most typical reason for hospitalisation) (Ingle et al., 2014; Shoaib et al., 2019). Patients frequently experience poor quality of life and also mental health problems such as depression (Chang et al., 2006; Cleland et al., 2012-13; Cohen et al., 2015).

Around 900,000 people in the UK suffer from HF (Cleland et al., 2012-13). More than 250,000 hospital discharges and deaths were recorded in England and Wales in 2006/2007 (Cleland et al., 2011). This number is set to rise due to an ageing population, improved survival rate and more effective treatment (NICE, 2010; Cleland et al., 2012-13). Unplanned readmissions are the main contributing factor to the financial burden on the NHS (Zhang et al., 2013).

The overall prognosis of patients with HF who require hospital admission is poor. One-year mortality of newly diagnosed patients is around 40% (NICE, 2010; Cleland et al., 2012-13). According to the National Heart Failure audit, there is much room for improvement in HF risk modelling (National Heart Failure Audit Team for, 2010; Allen et al., 2012). Increasing the number of HF medications alone may not be enough to reduce both the hospitalisations and annual mortality rates. (Cleland et al., 2011). However, many strategies to improve the outcome of patients with HF have been developed, and some are still in the process of development. Many research studies, including randomised controlled trials, suggest that intensified long-term management, together, where appropriate, with better patient education and effective monitoring of treatment, might substantially improve outcomes (Cleland et al., 2011; Cohen et al., 2015; Frohlich et al., 2019).

At present, when "risk" is estimated for individual patients, the risk being assessed is that of death or hospitalisation. These two are commonly used as end-points in clinical trials (Inglis et al., 2010). Electronic data offer a way of describing the trajectory of the disease course in groups of patients and individuals (Sutradhar et al., 2011; Poolsawad et al., 2014; Jiang et al., 2019). A model which describes how a group of patients might progress after an assessment for possible HF would be helpful to allow healthcare economies to understand how a whole

population of patients might behave (Braunwald, 2015; Krajewska et al., 2017; Zhang et al., 2018b; Kazmi et al., 2022).

1.2 Research Aims and Objectives

The overall aim of the study is to enhance our understanding of the dynamic nature of HF, which will, in turn, help us to develop risk models to predict and then *manage* risk better. The data source used in this thesis is the Hull LifeLab (HLL) (Masini et al., 2022), a large, epidemiologically representative population of patients with HF (see chapter 4). The thesis explores the potential methodology with a focus on HF, but is generic for chronic diseases, such as chronic lung or kidney disease. The study adopts a dual temporal perspective, focusing on both short term (immediate) and long term (extended) prediction. This allows for a comprehensive analysis of the disease's progression and improve risk prediction, thereby allowing the development of management strategies tailored to the needs of both patient groups and individuals. This may lead to the development more effective and timely interventions to improve patient outcomes.

The aims and objectives of the thesis are:

1. To explore the progression of HF using data.

   Investigate various methods for analysing and modelling changes in patients' health over time. This will allow for the development of a comprehensive understanding of how HF disease progresses in patients. Particularly the distinct phases of HF (such as, acute, unstable, and stable periods)

2. Investigate and explore the practical challenges present in clinical data for dynamic risk analysis of patient(s).

   There is need for both exploring and understanding the nature of data, and also structuring appropriately for any form of risk analysis. This will allow the development of appropriate strategies to improve handling of data, and thus make the construction of methodologies for assessing the risk of adverse events in HF patients in the long term. This will also allow for a better understanding of the definitions of HF health states and thus make modelling of risk at various stages more efficient.

3. Develop dynamic risk models and trajectories for both groups of patients and individual patients.

   Create advanced predictive models using machine learning techniques, based on Multi-state data. Identify the nature of the distinct phases of HF (such as, acute, unstable, and stable periods) from both the models and data. Thus allowing the demonstration of accurately tracking of health status changes and risk prediction for death or hospitalisation.

1.3 Thesis Structure

This thesis describes the various methodologies and algorithms for understanding the trajectories of HF. In order to do this, the thesis has 8 further chapters. These are as follows:

Chapter two: This chapter describes the nature of HF, and provides the context for understanding the progression of CHF (clinical perspective). There needs to be a better

understanding of difference between progressive and non-progressive chronic disease. Such an understanding will allow for the structuring of data for progressive risk analysis. This information is vital for implementing dynamic risk modelling, which incorporates current health states to predict future shifts in a patient's condition.

Chapter three: This chapter provides (methodological perspective) an overview of the dynamic risk methodologies. It explains the concepts of "trajectory" and "dynamic risk" analysis, and highlighting the key differences between the two approaches of risk modelling. The chapter also covers various methodologies and mathematical definitions used to represent changes in health status over time, highlighting the diverse approaches to modelling health trajectories.

Chapter four: This is where the data is described. It gives a detailed overview of HLL, outlining its structure and the clinical information it contains. For dynamic risk analysis and modelling, longitudinal data is required. However, the data in HLL consists of variables of patients at each visit to the clinics or hospitals; in other words, it is neither longitudinal nor episodic. Thus the data needs to be structured and reorganised into longitudinal health events. This will naturally involve all the challenges associated with the handling of clinical data, for example, misaligned samples; missing values, and non-normal distributions. As a result, the chapter offers detailed processes and recommendations for overcoming these issues, using specific examples that are generic across the board for these kinds of problems. Individual patient case examples are also used to both illustrate and enhance the understanding of disease progression in patients with HF. This chapter covers both clinical and methodological interplay.

In **Chapter five** machine learning techniques are applied to classify the HLL population. Two strategies are used to define the data, (a) Binary classification where the two classes are alive or dead and, b) where the alive patients are further divided. Thus the three classes are out-patient visit or hospitalisation and dead (multiclass classification). Three different machine learning methods are applied to the data to show how the models perform under different conditions. The chapter helps to identify data imbalances and extract other meaningful knowledge from a large dataset. It also illustrates the standard approach to dynamic risk modelling is naïve and perhaps insufficient for accurate predictions.

The development of dynamic risk models is carried out in **Chapter six.** Here the model is based on Absorbing Markov chains to model the progression of CHF. These models, require that the patients can be described using a finite number of mutually exclusive and exhaustive distinct states. The models allow us to see if events at an early stage can predict what is the likely eventual outcome during a subsequent follow-up. The data is then divided further using demographics of Age and Sex in order to better understand the differences in the progression of CHF and its dependencies on the demographics.

Chapter seven builds on the work in Chapter six by including clinical covariates (such as demography, aetiology, vital signs, blood test results, and treatment) to examine their influence on transitions between mutually exclusive clinical states. The scale of prediction also broadens, covering both short-term and long-term behaviours of the model. It also facilitates model validation over a five-year span based on yearly cycles. This allows us to derive the risk of transition between states, which in turn allows a greater understanding of the course of disease progression.

The results derived from the multistate risk model (developed in Chapter 7) are presented in **Chapter eight**.

The thesis is concluded in **Chapter nine**. In this chapter, the key findings are associated with the objectives, and they are discussed. Based on this the chapter also recommendations for future work, suggesting areas for further investigation and potential improvements to the methodologies developed in this study.

Chapter 2 Background

This chapter aims to build a foundation of understanding regarding the progression of CHF. This provides the essential background information and context form the clinical perspective needed for this thesis. The primary goal of this study is to predict risks in heart failure patients using a dynamic approach. By analysing patterns over time, I aim to enhance these patients' risk assessments' accuracy and timeliness. Understanding the dynamic nature of chronic diseases is crucial to illustrating how these conditions evolve and impact patients over time. It highlights the need for detailed approaches for effective management and intervention.

I started by exploring the complex nature of CHF, highlighting the challenges of managing long-term health conditions. The chapter also discusses the difference between progressive and non-progressive nature of chronic diseases to understand how their courses differ. Based on the foundational knowledge, I will explore various trajectories of heart failure can take in later chapters.

2.1 Dynamic nature of chronic diseases

Chronic diseases are long lasting health conditions that remain over a prolonged period, often spanning years or even a lifetime. (Megari, 2013; Diaz et al., 2015; Benkel et al., 2020). Chronic illness can be complex and usually have different stages, evolving due to various internal and external factors. Internal factors in the context of chronic diseases could include genetic predispositions, immune system function, hormonal imbalances, and overall health

status. External factors might encompass environmental exposures, lifestyle choices (such as diet, exercise, and stress levels), access to healthcare, socioeconomic factors, and exposure to infectious agents or toxins. These factors collectively contribute to the development and progression of chronic diseases.

Over the last decade, significant transformations have occurred in disease and health trend investigation. In traditional epidemiology the exploration of the association between exposure and outcome is often simplified into two states: having the disease or not having it (Shih et al., 2009). This approach limits the understanding of disease progression, especially for chronic ones (Pearce, 1996). Effective management of chronic diseases usually involves long-term treatment plans, including medication, lifestyle modifications (such as diet and exercise), and regular monitoring by healthcare professionals.

2.2 Progressive and non-progressive chronic diseases

Many chronic illnesses are progressive, meaning they worsen over time. Example includes certain types of cancer, Type II diabetes and Alzheimer's disease. Management focuses on slowing disease progression, controlling symptoms and improving quality of life. Chronic diseases with non-progressive nature may follow different trajectories, remaining stable or improving with treatment, and patients may experience periods of remission where symptoms are minimal or absent. Example include heart failure, asthma, rheumatoid arthritis and certain types of diabetes, all of which may exhibit different trajectories. There is a need for a comprehensive framework that helps to deal with the dynamic nature of non-progressive chronic diseases and addresses specific patterns observed during distinct phases, including acute, unstable, and stable periods (Corbin, 1998; Salonen et al., 2019; Naumzik et al., 2023).

2.3 The trajectory perspective in chronic disease

Implementing chronic illness trajectories in practice presents challenges (Hupcey et al., 2009), which often leads healthcare practitioners to prioritise symptom-focused care, perceived best practice over addressing the underlying patterns or courses of disease (Lubkin & Larsen, 2013). The practical application of chronic disease management requires longitudinal monitoring due to the recurrent and prolonged nature of trajectory phases, which can span weeks, months, or even years. A systematic framework (which encompasses the strategies, tools, or models) is needed to identify the shifts in the trajectory, providing valuable insights to health care professionals to adapt better and update treatment plans to address the evolving nature of the disease for more effective patient care (Moshkovich et al., 2020; Naumzik et al., 2023). This includes identifying recurrent phases and their relationship with symptoms and knowing when to update the treatment plan, among other considerations.

2.3.1 Key elements of health trajectories research

The concept of health trajectories refers to the patterns over time. The understanding the dynamics and reasons behind these changes is crucial for identifying those at greatest risk of adverse events (Henly et al., 2011; Salonen et al., 2019). Health trajectory research focuses on understanding changes in individuals, groups, or communities. This contrasts with cross-sectional research which examines differences among individuals at a single time point. While findings from typical cross-sectional designs are focused on variable-centred and remain static (Salonen et al., 2019). Health trajectory research includes various approaches (Henly et al., 2011):

- Observational studies examine natural health progression over time.
- Experimental studies use trajectories to understand intervention effects.
- Clinical course investigations analyse individuals' clinical journey, considering self-care and professional interventions.

2.3.2 Longitudinal data

Longitudinal data involves collecting information from individuals or groups over time. This type of data is essential for studying dynamic nature of chronic diseases (Sutradhar et al., 2011; Poolsawad et al., 2014; Sutradhar & Barbera, 2014; Jiang et al., 2019). Statistical analysis using longitudinal data provides insights into the natural progression of health over time, facilitates the identification of patterns and trends (Roger, 2013; Braunwald, 2015; Krajewska et al., 2017; Zhang et al., 2018a). This allow researchers to investigate adverse events, enables them to development of predicted models for different patient groups or at population levels. Also, help design personalised interventions to address patients' unique needs effectively.

In longitudinal data analysis, trajectories are characterised by several key components: an initial point (the value at a specified time zero), a form or shape defined by a mathematical function, the rate of change over time (speed), and any changes in the rate (acceleration) (Henly et al., 2011).

To simplify the above statement, I can say that trajectories are like paths that show how things change over time. They start from a certain point, follow a particular shape or pattern, and can speed up or slow down (rate of change) as time passes. Sometimes, they might even change how fast they change (acceleration). These descriptions help us understand the

dynamics of changes (how things develop or evolve throughout a study). In trajectory analysis, considering the timing ("time") and circumstances ("context") surrounding changes in health status or outcomes become more crucial for accurately capturing and interpreting patterns of development or decline (Henly et al., 2011; Neale, 2015). Incorporating these two dimensions into research methodologies enhances the depth and applicability of findings, providing a more comprehensive understanding of how chronic diseases evolve and how individuals' health trajectories are influenced by their life circumstances.

2.3.3 Classification of patients into subgroups

The health trajectory modelling aims to classify individuals into distinct subgroups who share similar response patterns. These patterns can vary significantly depending on the nature of the study and the variables being measured. For example, patients may be grouped based on different trajectories of symptom severity, such as variations in pain intensity scores over time (Busch, 2002). In a study tracking the effectiveness of a new medication for pain management, the response pattern might show how pain levels fluctuate over time in participants receiving the medication compared to those receiving a placebo.

After identifying subgroups the trajectory classes can serve as either a dependent variable, helping to identify predictors of health trajectories, or an explanatory variable, allowing exploration of their effect on future health outcomes (Nguena Nguefack et al., 2020). Trajectory membership refers to categorising individuals into specific groups or subgroups based on their patterns of change over time about certain variables or outcomes. Figure 2.1 illustrates the trajectory of rectal temperature for the first patient who underwent direct visualisation of their open-heart surgery. The procedure performed at the University of

Minnesota in 1952. This was honoured on the Wall of Discovery along the campus Scholars Walk. (Bolman & Black, 2003).

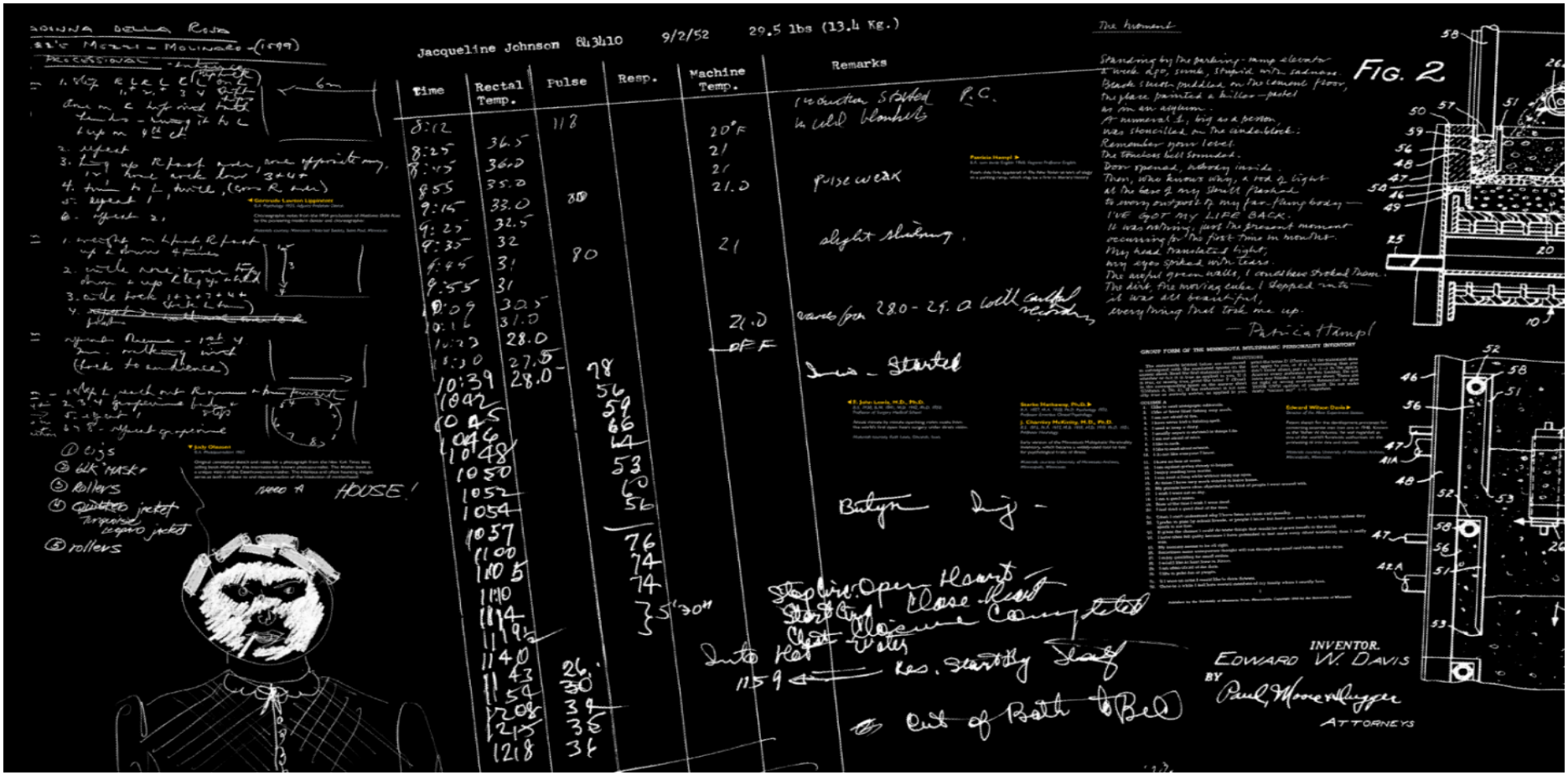


Figure 2.1: Direct visualisation of first open-heart surgery.

This figure is taken from the study (Henly et al., 2011). The surgery employed a simple hypothermia method, using refrigerated blankets to decrease the temperature and stop the heart, creating a bloodless surgical field. The temperature was expected to follow curvilinear path: starting from normal, it decreased to and was maintained at 28°C (82.4°F) during the operation, and then re-warmed in a water bath until reaching 36°C (96.8°F). (Photo courtesy of LA ink)

Figure 2.2 provides a more detailed understanding of how health outcomes change over time compared to methods solely based on average values from a sample. The example illustrates the method for modelling pain intensity: population average models (top) and trajectory modelling approach (bottom). Both models utilize longitudinal data from the same patients collected at four occasions (T0, T1, T2, T3 (equally time spaced between each point)). In the top graph, the mean score was calculated at each time point for the entire population. In the bottom graph, researchers characterized subgroups based on pain intensity levels and calculated mean scores based on trajectory membership levels.

This approach enhances understanding of intra- and inter-individual variability and patterns of health outcomes over time. It is valuable for examining the diversity of health profiles, identifying vulnerable subgroups needing improved healthcare, and pinpointing trajectories leading to the best health outcomes.

- Intra-individual variability refers to the variations or fluctuations in health outcomes observed within the same individual over time. For example, a person's blood pressure may fluctuate throughout the day or over a week.
- Inter-individual variability: This refers to the differences in health outcomes observed between individuals within a population at any given time. For example, one person may have higher blood pressure than another person at the same age and with similar lifestyle factors.

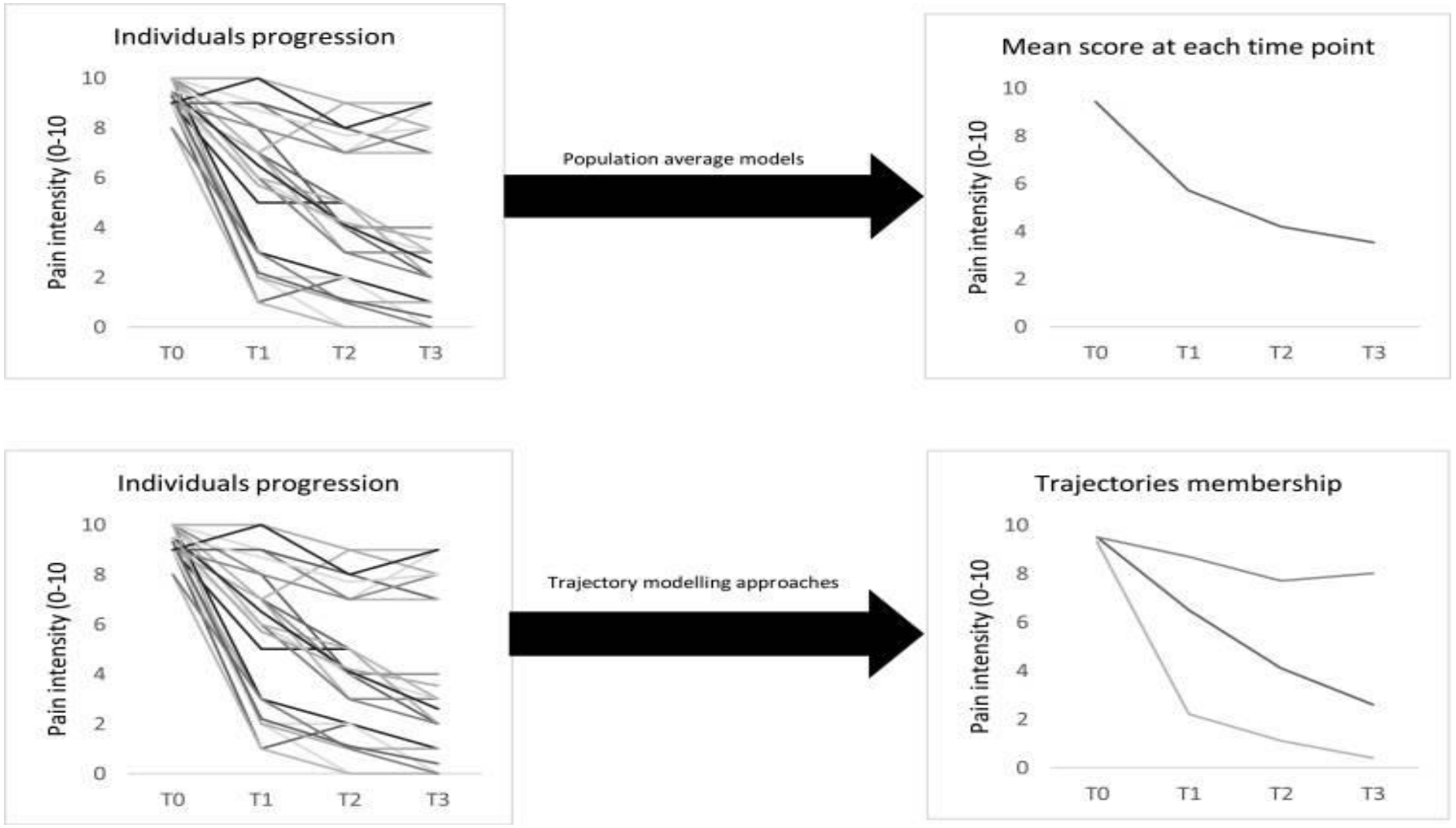


Figure 2.2: Methods for modelling pain intensity taken from the (Nguena Nguefack et al., 2020).Illustrated as population average models (top)

Figure 2.3 depicts two sets of trajectories: pre-test and potential outcome trajectories after an intervention, with time represented as "T". The study (Campbell & Stanley, 1963) highlights the significance of observing changes over time to understand the effects of a treatment or intervention. By comparing the trajectories before and after the intervention, researchers can draw meaningful conclusions about how the treatment affects the outcomes of interest. This before-and-after comparison is crucial for evaluating the effectiveness of interventions and drawing insights into their impact on the studied variables.

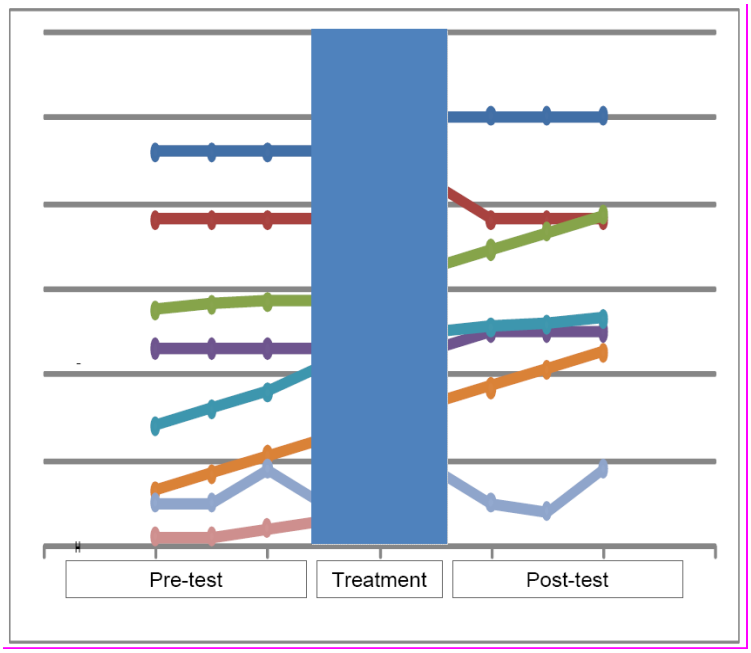


Figure 2.3: Pre-test and post-test trajectories.

This diagram illustrates individual changes before and after an intervention, as shown in the shaded area. The idea is taken from Campbell's "Experimental and Quasi-Experimental Designs for research", 1st edition, ©1966 Wadsworth, a part of Cengage Learning, Inc. credit to author.

2.3.4 Time-dependent data

In health services, understanding patterns in health status during acute illness relies heavily on time-dependent data (Lyons et al., 2023). This includes health history obtained through interviews or written self-reports, stored in either paper format or electronic databases within medical records. Such data serves as a critical mechanism for diagnosing various illnesses and developing effective treatment plans aimed at optimizing health over time (Nguena Nguefack et al., 2020). This emphasizes the importance of considering time-dependent factors in healthcare to provide timely and appropriate interventions tailored to individual needs.

2.3.5 Change and Time

Change, defined as the alteration of a state or experience, is a fundamental aspect observed naturally and intentionally in research and clinical settings (Collins & Horn, 1991). Traditional methods of describing change, such as subtracting scores from different occasions, have known limitations (Harris & Youth, 1967), prompting the need for newer approaches focusing on modelling data over multiple occasions (Singer & Willett, 2003; Fairclough, 2010). Using modern statistical techniques enable researchers to capture and describe the patterns of change within each person's (intra-individual) health outcomes and clinical variables over time (Nesselroade & Ram, 2004; Nguena Nguefack et al., 2020). (Kazmi et al., 2022) suggests that in heart failure disease research, the trajectory analysis (baseline, rate and direction of change) offers better predictive value for functional outcomes than simply comparing the mean difference between two groups.

The term 'parameters' in the study refers to the coefficients, constants, equations or functions used in Markov and multi-state modelling to represent relationships between variables. These parameters are estimated from the data and can provide insight into the nature and strength of relationships between variables.

Time is crucial in understanding the evolution and interaction of health-related variables. Accurately defining and conceptualizing time intervals for trajectory analysis presents challenges (McGrath & Tschan, 2004). The relationship between time, individual health and illness can exhibit both discontinuous and continuous patterns. Continuous change may follow smooth and predictable trajectories, described by simple mathematical functions, or fluctuate in complex ways with multiple fluctuations over time. Such variability adds complexity to understanding how individuals' health changes.

Time can moderate the relationship between covariates and changing health status, such as age. This process requires a well-designed measurement plan and a thoughtful approach to measuring and representing time by the specific research question (Singer & Willett, 2003). Researchers should carefully select the type of time measurement, whether using clock time (seconds, minutes, hours), calendar time (days, weeks, months, years), biological time (natural internal rhythms that regulate physiological processes), or social time (Henly et al., 2003).

2.3.6 Design

Issues as follows:

- Developing a measurement protocol: This involves creating a systematic plan for collecting data to uncover trajectories of change in health indicators over time.
- Establishing a time measurement and coding scheme: This step involves defining how time will be measured and coded in the study, ensuring consistency and accuracy in tracking the progression of health trajectories.
- Utilising appropriate instrumentation: Researchers must select suitable tools and methods for effectively collecting and analysing longitudinal data.
- Cautious selection of the scale and coding of time: The choice of scale (metric) and coding method for time measurement is crucial for accurately depicting the nature of the health trajectory (continuous or discontinuous) and interpreting its progression.

Since there is no standard starting point for time, researchers must decide when it begins (initial time points or baseline (T0)) for a trajectory, typically selecting theoretically significant time-points like the child's birth, hospital admission or discharge, or the diagnosis of a chronic or terminal illnesses. When conducting a study or research, it's important to carefully determine how long the observations will take place (the observation period) and how often these observations will occur (the frequency of observations). More complex change models may require more timed observations to capture key features accurately. Individuals may undergo observations at varying schedules, or time-points, and on different occasions within the designated trajectory timeframe (described in later chapters) (Singer & Willett, 2003).

2.4 **Conclusion**

This chapter has provided the necessary background and context to understand the progression of chronic illness, especially chronic heart failure. It also highlighted the need for detailed management strategies and interventions, further explored in the next chapter. The chapter differentiated between progressive and non-progressive diseases, setting the stage for subsequent discussions on the varied trajectories heart failure can take, which will be further explored in later chapters. The importance of longitudinal data in tracking disease progression and distinguishing between patients' current and previous health states is also highlighted. This understanding is vital for implementing dynamic risk modelling, incorporating current health states to predict future patient conditions shifts.

Chapter 3 CHF: Trajectory analysis and Risk modelling

To uncover the underlining progression pattern of the CHF requires more advanced predictive models (Braunwald, 2015; Krajewska et al., 2017; Zhang et al., 2018a). Standard statistical tools often lack the dynamic adaptability and precision that artificial intelligence (AI) and machine learning (ML) can provide. Given the poor overall prognosis for patients with heart failure who require hospital admission (Shoaib et al., 2019; Abel et al., 2024) and the high mortality rate (NICE, 2010; Cleland et al., 2012-13; Sokoreli et al., 2016; Shiraishi et al., 2018), there is significant room for improvement in HF risk modelling, as highlighted in chapter 2.

The chapter highlights the diverse approaches used to achieve the thesis's aim of dynamically predicting risk in patients with HF. This requires a deep understanding of the concepts of "trajectory" and "dynamic risk" analysis. The concepts for these two methods are related yet different from each other. Understanding both concept is essential for developing risk prediction models that aid in patient care and healthcare planning for CHF. Various methodologies and mathematical definitions used to represent changes in health status over time are discussed in later sections.

3.1 Statistical models for change

As discussed above, for patients with CHF, the clinical interest encompasses the dynamics of disease progression, not just the final outcome (Khand et al., 2001; Ieva et al., 2017; Jiang et al., 2019; Kazmi et al., 2022). Longitudinal data with time-dependent capability

offer ways to describe the change in the trajectory of the disease course across various groups of patients (Sutradhar et al., 2011; Poolsawad et al., 2014; Jiang et al., 2019). Constructing a model to describe how a group of patients might progress following an assessment for potential heart failure could be beneficial. (Braunwald, 2015; Krajewska et al., 2017; Zhang et al., 2018a).

I aim to model the progression of CHF through its acute, unstable, and stable phases by translating the trajectory framework into a data-driven dynamic model. Changes in health trajectories can manifest in various ways, including stability (no change), gradual improvement or deterioration, acceleration or deceleration of progression, or complex fluctuations across different phases. Capturing these details requires sophisticated statistical techniques. Addressing these challenges allow researchers can gain deeper insights into the progression of chronic conditions like heart failure and develop more tailored interventions to improve patient outcomes (Cudeck & Klebe, 2002; Collins, 2006).

3.2 Approaches used for analysis and modelling change

Three main methods are used to study change patterns or sample subgroups.

Nonparametric methods, such as clustering algorithms like k-means and hierarchical clustering, do not assume any specific data distribution. These techniques rely on dissimilarity measures to assign individuals to subgroups, making them valuable for grouping individuals based on similarity without preconceived assumptions about data generation. Here's an example to illustrate the concept: let's say I have a dataset containing information about the health status of individuals over time, including variables like blood pressure, cholesterol levels,

and body mass index (BMI). Based on these variables, I want to identify subgroups of individuals with similar health trajectories. *k-means clustering*, the algorithm groups individuals into clusters based on their similarity in health status over time. It does this without assuming any specific distribution of the data. The algorithm iteratively assigns each individual to the cluster whose mean health status is closest to their observed values. *Hierarchical clustering* groups individuals based on the similarity of their health status trajectories, forming clusters hierarchically. At each step, the algorithm merges the two most similar clusters until all individuals are grouped into one large cluster. These nonparametric methods allow us to identify meaningful subgroups of individuals with similar health trajectories without imposing any assumptions about how the data were distributed.

Parametric approaches, like Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs), presuppose that data arise from finite mixtures of distributions, assigning subgroups based on conditional probabilities. Let's consider an example in a clinical setting where I am analysing trajectories of depression severity over time among patients undergoing treatment. *Gaussian Mixture Model - GMM:* In this approach, I assume that the distribution of depression severity scores among patients follows a mixture of Gaussian distributions. Each Gaussian component represents a subgroup of patients with distinct trajectories of depression severity. For instance, one component might represent patients who experience a rapid reduction in depression severity after starting treatment. In contrast, another component might represent patients whose severity remains relatively stable over time. The assignment of patients to these subgroups is determined based on the conditional probabilities of their observed depression severity scores given the parameters of the Gaussian mixture model. *Hidden Markov Model - HMM:* Here, I assume that patients transition between different states of depression severity, with each state corresponding to a different trajectory pattern. A set of hidden transition probabilities governs the transitions between states. These transitions are not

directly observable but are inferred based on the observed depression severity scores. Patients' subgroup assignments are determined based on the sequence of observed depression severity scores and the estimated parameters of the HMM.

Semi-parametric approaches combine features of both nonparametric and parametric methods. Doing so offers a versatile modelling approach while still integrating certain parametric assumptions. This balance enables researchers to capture complexity in trajectory analysis while benefiting from the structure provided by parametric models. Examples of semi-parametric methods include Local Polynomial Fitting and Kernel Density Estimation. This categorization offers a structured overview of available methods for analysing trajectory patterns and identifying sample subgroups, aiding researchers in selecting the most suitable approach for their studies. Let's consider a clinical study investigating the trajectories of pain intensity following a surgical procedure. In this scenario, semi-parametric approaches could be applied to analyse the trajectories of pain intensity over time. For instance, researchers may use: *Local Polynomial Fitting* to model the trajectory of pain intensity, allowing for flexibility in capturing complex patterns of change while still incorporating some parametric assumptions. This method could help identify subtle variations in pain intensity trajectories among different patient subgroups, such as those with varying preoperative pain levels or different surgical interventions. *Kernel Density Estimation* could be employed to estimate the probability density function of pain intensity at different time points post-surgery. This approach would provide a smooth representation of the distribution of pain intensity over time, allowing for identifying distinct peaks or clusters in the trajectory data.

By utilising semi-parametric methods like Local Polynomial Fitting and Kernel Density Estimation, researchers can effectively analyse the trajectories of pain intensity while striking a balance between flexibility and structure in their modelling approach. This enables them to uncover nuanced patterns in pain intensity trajectories and identify subgroups of patients with differing experiences of post-surgical pain.

3.3 Functions

Figure 3.1[1] displays various mathematical functions used to represent changes in health status over time (Cudeck & Klebe, 2002). Each function has a unique set of rules, known as parameters, which create distinct graph patterns.

- Constant functions maintain the same value over time, but they can be distinguished by a parameter called κ, which determines their starting level when time is 0.
- Linear functions represent steady changes and are defined by two parameters: π0 represents the initial level at time 0, and π1 represents the rate of change (often referred as the slope) per unit time.
- Quadratic and higher-order polynomial functions effectively illustrate changes that accelerate or decelerate over time.

[1] H(t) denotes health over time, with constant, linear, and quadratic functions being types of polynomials. The exponential function uses the constant e, approximately 2.718, which is serve as base of natural logarithms. A piecewise function changes its form at a specific point in time t_j, where j indicates the change point; usually, graphs of piecewise functions are linear on either side of t_j.

.

- Exponential functions exhibit extreme values at the beginning (ξ), eventually levelling off (α, the asymptote), and vary in the time it takes to transition from the extreme value to the levelled-off value (ρ, the rate).
- Piecewise functions represent different types of changes during distinct periods.

These mathematical functions offer valuable tools for visualizing and understanding the dynamics of health status changes over time.

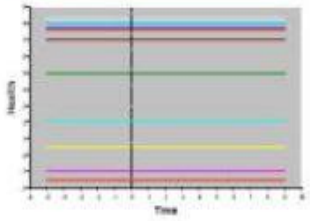
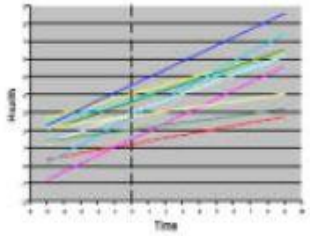
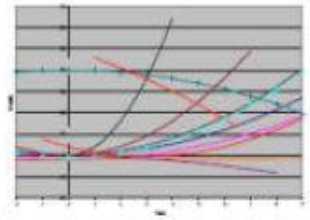
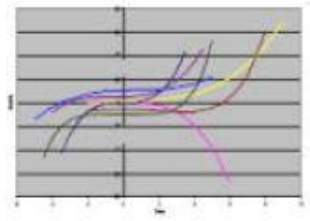
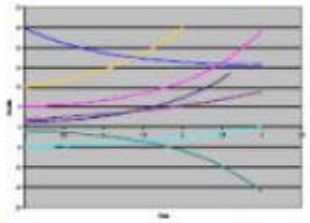
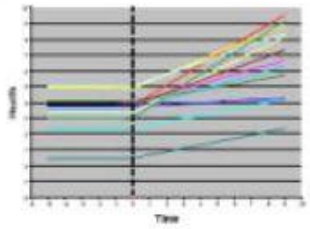
Function	Parameters	Typical Graphs
Constant $H(t) = \kappa$	κ = y-intercept	
Linear $H(t) = \pi_0 + \pi_1 t$	π_0 = y-intercept π_1 = slope (rate of change)	
Quadratic $H(t) = \pi_0 + \pi_1 t + \pi_2 t^2$	π_0 = y-intercept π_1 = linear change coefficient π_1 = quadratic change term	
Higher-Order Polynomial $H(t) = \pi_0 + \pi_1 t + \pi_2 t^2 + \ldots + \pi_n t^n$	π_0 = y-intercept π_1 = linear change coefficient π_1 = quadratic change coefficient π_n = n^{th} degree change coefficient	
Exponential $H(t) = \alpha + (\xi - \alpha)e^{\rho t}$	α = asymptote (leveling out value) ξ = extreme value at t = 0 ρ = rate of change between ξ and α	
Piece-wise $H(t) = H_1(t), t < t_j$ $H(t) = H_1(t), t \geq t_j$	Parameters of H_1 Parameters of H_2	

Figure 3.1: Mathematical function used to model change (Henly et al., 2011)

3.3.1 Complementary techniques and extensions

The advancements in statistical modelling over the past three decades have led to the development of more flexible versions of the basic model for studying changes (Skrondal & Rabe-Hesketh, 2004). These extensions, as summarised in Figure 3.2, serve various purposes and offer researchers various options depending on their research questions and the data available. For modelling changes in health outcomes measured in longitudinal studies, researchers can choose from various approaches categorized into three main groups: nonparametric, parametric, and semi-parametric methods and algorithms (described above). Each of these approaches has its strengths and limitations. (Nguena Nguefack et al., 2020).

Model/Approach	Purpose	Example
Time-varying Covariates	To estimate the effects of an individual's changing status on trajectory parameters	Harrison, T., Blozis, S., & Stuifbergen, A. (2008). Longitudinal predictors of attitudes towards aging among women with multiple sclerosis. *Psychology & Aging, 23*, 823-832. doi: 10.1037/a0013802
Trajectories in Context	To estimate the effects of contextual factors on trajectory parameters	Black, M. M., & Krishnakuman, A. (1999). Predicting longitudinal growth curves of height and weight using ecological factors for children with and without early growth deficiency. *Journal of Nutrition, 129*, 539-543.
Latent Class Growth Model	To identify latent classes based on similar trajectories, without variation within class	Gill, T. M., Gahbauer, E. A., Han, L., & Allore, H. G. (2010). Trajectories of disability in the last year of life. *New England Journal of Medicine, 362*, 1173-1180.
Growth Mixture Model	To identify latent classes based on similar trajectories, allowing individual variation within class	Stoddard, S. A., Henly, S. J., Sieving, R. E., & Bolland, J. (2010). Social connections, trajectories of hopelessness and serious violence in impoverished urban youth. *Journal of Youth and Adolescence*. doi: 10.1007/s10964-010-9580-z
Parallel Process Model	To estimate multiple trajectories and relationships among growth parameters for each trajectory	Taylor, M. G., & Lynch, S. M. (2004). Trajectories of impairment, social support, and depressive symptoms in later life. *Journal of Gerontology: Social Sciences, 59B*, S238-S246. doi: 10.1093/geronb/59.4.S238
Categorical Responses	To estimate trajectories based on categorical responses by incorporating a link function into the model	Hedeker, D., & Mermelston, R. J. (2000). Analysis of longitudinal substance use outcomes using ordinal random-effects regression models. *Addiction, 95 (Supp. 3)*, S381-S394. doi: 10.1080.09652140020004296
Generalized Estimating Equations	To estimate a population average model for longitudinal data using a semiparametric regression approach	Bohl, A. A., Fishman, P. A., Ciol, M. A., Williams, B., LoGerfo, J., & Phelas, E. A. (2010). A longitudinal analysis of total 3-year healthcare costs for older adults who experience a fall requiring medical care. *Journal of the American Geriatrics Society, 58*, 853-860. doi: 10.1111/j.1532-5415.2010.02816.x
Dynamic Process Models	To estimate parameters of dynamic systems using coupled differential equations models	Boker, S. M., & Laurenceau, J.-P. (2006). Dynamical systems modeling: An application to the regulation of intimacy and disclosure in marriage. In T. A. Walls & J.L. Schafer, *Models for intensive longitudinal data* (pp. 195-218). Oxford: University Press.

Figure 3.2: Enhancements of the standard random-effect model for change. Supplementary methods and case studies (Henly et al., 2011)

3.4 Illness trajectory vs dynamic risk analysis for heart failure

As explained earlier, the concepts of "trajectory" and "dynamic risk" modelling are related but distinct in terms of definition and purpose. However, in this thesis, I used trajectory to understand the patterns in disease and based on these changing patterns I then developed the dynamic risk models.

3.4.1 Trajectory analyses

- **Definition:** A trajectory in healthcare refers to the path or progression of a disease or health condition over time. This can include the rate at which a disease progresses, its stages, and any potential outcomes.
- **Purpose:** The concept of trajectory is used to categorize and predict the course of a disease in an individual or group. Understanding these trajectories helps healthcare providers anticipate patients' needs, manage expectations, and plan interventions.
- **Focus:** Trajectories are primarily concerned with the pattern of a disease's progression or stability. They provide a longitudinal view of how a condition evolves, highlighting improvement, stability, or decline.

3.4.2 Dynamic Risk analysis

- **Definition:** Dynamic risk modelling involves using statistical techniques and data analysis to predict future health outcomes based on current and historical patient data.

- **Purpose:** The main goal of dynamic risk modelling is to identify individuals at high risk of developing a disease or experiencing a worsening of their health condition. These models enable targeted interventions, efficient allocation of resources, and personalised care plans to mitigate risks.
- **Focus:** Unlike trajectories, which describe the course of a disease, dynamic risk models are focused on quantifying the likelihood of future events or outcomes. They integrate various risk factors and patient data to provide a probabilistic assessment of what might happen to a patient's health status in the future.

3.4.3 Key differences

- **Application:** Trajectory analysis is used to understand and describe the course of a disease, while dynamic risk modelling is used to predict future health outcomes and inform intervention strategies.
- **Data Use:** Trajectories are mapped based on longitudinal observations of disease progression. Dynamic risk models utilize these observations along with a broader range of data inputs (e.g., demographic information, lifestyle factors, genetic information) to generate predictions.
- **Outcome:** The outcome of trajectory analysis is a descriptive understanding of how a disease progresses. In contrast, the outcome of dynamic risk modelling is a predictive insight into individual or population-level risks.

3.5 Dynamic process models

My main focus is dynamic process modelling, which will offer a comprehensive approach to a) capture changing needs of patients with heart failure, b) foster more nuanced and effective management strategies and c) improve patient quality of life (QoL), decrease the burden of hospitalizations, and extend survival.

Patient will be categorised according to identified patterns and stages of the HF disease. These stages can be categorised broadly into:

- Stable: Patients may have periods where symptoms are controlled, and the disease progresses slowly.
- Acute Decompensation: Episodes where symptoms suddenly worsen, often requiring hospitalization.
- Gradual Decline: A steady worsening of the condition with increasing symptoms, frequent hospitalizations, and reduced quality of life.
- Terminal: These are the final stages of heart failure, when treatments become less effective, and palliative care may be considered.

These categories of HF patients will be used to forecast the likelihood of future adverse events, such as hospital readmission, period of stability or deterioration in health status, or mortality. Key components of dynamic risk modelling include:

- **Data integration:** Incorporating a wide range of data, including clinical measurements, patient-reported outcomes, laboratory results, and more.

- **Risk Stratification:** Classifying patients based on their risk of future adverse events, which can guide the intensity of monitoring, treatment decisions, and resource allocation.
- **Predictive modelling:** Employing statistical and machine learning techniques to analyse patterns in the data and predict future outcomes.

Trajectory analysis is typically performed with discrete-time data (Salonen et al., 2019) and dynamic process models are more versatile in handling discrete and continuous data and use differential equations. Based on the nature of the underlying processes, dynamic process models can be categorized into two main types.

1. **Discrete Stochastic Processes:** In this dynamic process model, the state variables change in discrete steps or intervals.
2. **Continuous Stochastic Processes:** These models involve systems where changes occur continuously over time.

The choice between discrete and continuous stochastic processes depends on the modelled system's nature and research objectives. It's important to note that growth functions are more commonly associated with continuous processes, while Markov chains are typically used to model discrete-time processes. The use of growth functions within the context of Markov chains may require additional considerations and adaptations to fit the discrete-time nature of the model (implementation of such format is discussed in chapter 6).

3.5.1 Markov chains models

Markov chains offer valuable insights into the progression of CHF, but the effectiveness of these methods is hindered by challenges in organising and analysing longitudinal data. Longitudinal data, obtained by following individuals over extended periods, often lack proper organisation, making it difficult to extract meaningful insights into disease progression. Additionally, there is a notable absence of comprehensive guides that cover the basics, data types, procedural steps, and suitable statistical tools for analysing CHF progression. Furthermore, the lack of clear instructions for reporting the results of such modelling exacerbates the problem, hindering the dissemination of findings and limiting their utility in clinical practice.

Markov chains model a wide range of dynamic processes in healthcare research by providing a flexible framework for analysing and simulating transitions between different health states over time (Andersen & Keiding, 2002). These models are valuable tools for understanding disease progression, treatment effectiveness, and predicting future health outcomes. Many chronic diseases, such as diabetes, cancer, and cardiovascular (CV) diseases, involve a progression through different health states over time. Markov chain models can represent these transitions probabilistically, where each health state represents a specific stage of the disease (e.g., asymptomatic, symptomatic, severe complications) and transitions occur between these states based on probabilities.

In the context of Markov chains, the differential equation typically used is the *Chapman-Kolmogorov equation*, which describes the evolution of the transition probabilities over time. Mathematically, it can be expressed as:

$$P_{ij}^{(n+m)} = \sum_k P_{ik}^{(n)} P_{kj}^{(m)}$$

Where:

- $P_{ij}^{(n+m)}$ represents the probability of transitions from state i to state j in $n + m$ steps.
- $P_{ik}^{(n)}$ represents the probability of transitioning from state i to state j in n steps.
- $P_{kj}^{(m)}$ represents the probability of transitioning from state k to state j in m steps.

This equation essentially indicates that the likelihood of transitioning from state i to state j in $n + m$ steps is obtained by summing the probabilities of transitioning from state i to any intermediate state k in n steps, and then from state k to state j in m steps. Chapters 6 and 7 cover further details of Markov chains in the development phases of the dynamic risk model for HF.

3.6 **Conclusion**

The chapter has provided an overview of dynamic risk methodologies, focusing on the key concepts of "trajectory" and "dynamic risk" analysis. By explaining these concepts, I have highlighted the fundamental differences between the two approaches and illustrated how each can be applied to understand and predict changes in health status over time. The chapter detailed various approaches used to model the changes, emphasising the importance of capturing the dynamic nature of disease progression.

Insight gained from chapter 2 and 3 sets the stage for the subsequent chapters where I translates the concepts of change into multistate models that provide meaningful insights and model the changes in health status of patients with heart failure. Next chapter will explore different 'risk states'— that describe the levels of severity or stability at any given time of patients with heart failure. These risk states help us understand where a person is in their journey with the disease. The time intervals between clinical examination and other health events is also important (whether it is days, months, or years) to reveal the rate at which the heart failure progresses.

In summary, while disease trajectories provide a framework for understanding the progression of health conditions over time, dynamic process models aim to capture the dynamic variability of chronic conditions by considering individual differences and temporal evolution perspectives. This means considering the changes in a patient's health status from a time-based viewpoint, taking into account how these changes occur over short and long periods. Markov chains will be utilised in the upcoming chapters to forecast the future shifts in a patient's condition based on their current health state. I also highlighted the importance of longitudinal data which is key in tracking disease progression. It helps distinguish between patients' current and previous health states.

Chapter 4 Data and methods

This chapter builds on previous discussions in chapters 1, 2 and 3 about CHF and risk modelling. The following sections of this chapter will provide a detailed overview of HLL, outlining its structure and the clinical information it contains. It will also discuss the challenges associated with the handling of clinical data and offers detailed processes and recommendations for overcoming these issues, using specific examples that are generic across the board for these kinds of problems. Individual patient case examples are also used to both illustrate and enhance the understanding of disease progression in patients with HF.

It has been explained in the chapter 1 that much of the understanding of the natural history of HF comes from clinical trials and epidemiological studies. However, clinical trials recruit highly selected populations, followed for a short period, to answer specific limited questions only. Epidemiological studies do not usually describe detailed patient information and are often not longitudinal. Partly to understand the apparent discrepancies, the HLL was designed to help understand the modern epidemiology of HF. The purpose is to recruit a large, epidemiologically representative population of patients with HF that bridges the gap between clinical trials and clinical practice. Utilising longitudinal data obtained from HF patients, I aim to enhance my understanding of the dynamic nature of heart failure, which will help use develop risk models to predict and manage the risk trajectory better.

4.1 Data collection method

4.1.1 Hull and East Riding of Yorkshire

Hull is unusual in the United Kingdom in being relatively geographically isolated. The population is stable with little emigration or immigration. The single Hospital Trust provides secondary care to a population of around 500,000 patients. Very few patients are admitted to hospitals elsewhere as there are no others within easy travelling distance for the majority of people. Hull is deprived but also provides care to the East Riding of Yorkshire population, a much more affluent area.

4.1.2 Referral of patients

A single community heart failure clinic was established by the academic department of cardiology in 1998. Referrals are accepted into the service from primary care, as well as from cardiologists and other secondary care physicians. Case-finding of in-patients allows referral of all recently hospitalised patients to the service. The service aims to see patients within two weeks of referral to comply with National Institute for Health and Clinical Excellence (NICE) guidance. With each appointment sent out, the patient receives information about the service, which also describes the purpose of HLL, together with a quality of life questionnaire.

Some patients have not been previously diagnosed with heart failure, require the initiation of guideline-recommended therapy. Others with pre-existing diagnosis are already on treatment that may need optimisation.

4.1.3 Structure of service

When first seen in the service, regardless of eventual diagnosis, patients are asked to sign a consent form which allows their data to be used in the study. Only those patients who consent to inclusion in HLL are included in reports. I cannot know (by definition) how many patients have declined to be included.

Each patient is then seen by a doctor who is either a trainee or a consultant cardiologist, with a heart failure specialisation. The patient is systematically reviewed and examined; then has a 12 lead electrocardiogram; chest x-ray, and detailed echocardiogram by an echo cardiographer with accreditation from the British Society for Echocardiography (BSE) following BSE guidelines; spirometry, and standard blood tests, including N-terminal pro-B-type natriuretic peptide. Blood is also taken and immediately centrifuged for storage as plasma and serum in a dedicated -80º C freezer.

At the initial visit (baseline (BL)), a diagnosis is made, and a treatment plan is initiated. The service is designed principally for patients with HF due to reduced left ventricular ejection fraction. Still, there are established pathways for patients thought to have valvular heart disease, a primary arrhythmia, or symptomatic coronary heart disease. Patients with HF and normal ejection fraction are also followed. Patients who have no cardiac diagnosis are discharged to their primary care physician.

Initially, the service was structured that patients were seen four months after their initial diagnosis, but as demand has grown, most patients are now seen at six-month intervals. A nurse-led service in the community allows up-titration of guideline-directed therapies between visits and provides an early review for more vulnerable patients.

4.1.4 Data collection

Data are collected using an Access® database. Data entry is as far as possible digital with minimal use of text fields. It features a relational database structure, containing a main (central) table with unique identifier for each patient, ensuring consistent data recording across other tables. There are tables related to past medical history, including co-morbidities. There are linked tables collecting separate data on the core modules: physical examination, drug history, electrocardiogram, echocardiogram, quality of life, spirometry, and blood tests. At each visit for each patient, an entry is made into each of the core tables. There are also tables collecting information as required on device implantation, cardiac MRI and angiography results and 6 minute walk test distance. This information is not need to be collected for all patients. A particular effort is made to ensure that each patient in the database has a minimum baseline dataset consisting of history, physical examination, medication, NTproBNP, ECG and echocardiogram. However, routine clinical testing for NTproBNP only became available from around mid-2003. The database also has a decision support tool: advice based on blood pressure, heart rate and rhythm, ECG, and current medication prompts the treating physician to consider the next therapies. A concise report is generated for the primary care physician and hospital records for each clinic visit. The report summarises diagnosis and treatment recommendations made by the physician seeing the patient.

The data are linked to the Office for National Statistic (ONS) mortality data to get the date and cause of death. I used the hospital episode statistics (HES) to determine hospital admissions at each 4-month interval, using data from 2000 to 2017.

A significant portion of clinical research focuses on developing clinical risk models using logistic regression or Cox proportional hazards (PH) regression. These models typically don't need long-term data; instead, they often use just the baseline clinical data or maximum data of two time-points (baseline and 1 follow-up) to analyse outcomes like death or hospital readmission. When developing a dynamic model for predicting the progression of heart failure disease, the longitudinal data is needed. HLL presents valuable opportunities of identification of risk factors and diseases progression over time. Individual patient case vignettes are presented to better understand the complexities involved. Figure 4.1 graphically represents each patient's healthcare journey, capturing key events such as OPD follow-ups, admissions, and death over time. This visual demonstrates how the data is organized using different methods and tools (as described in later sections of this chapter).

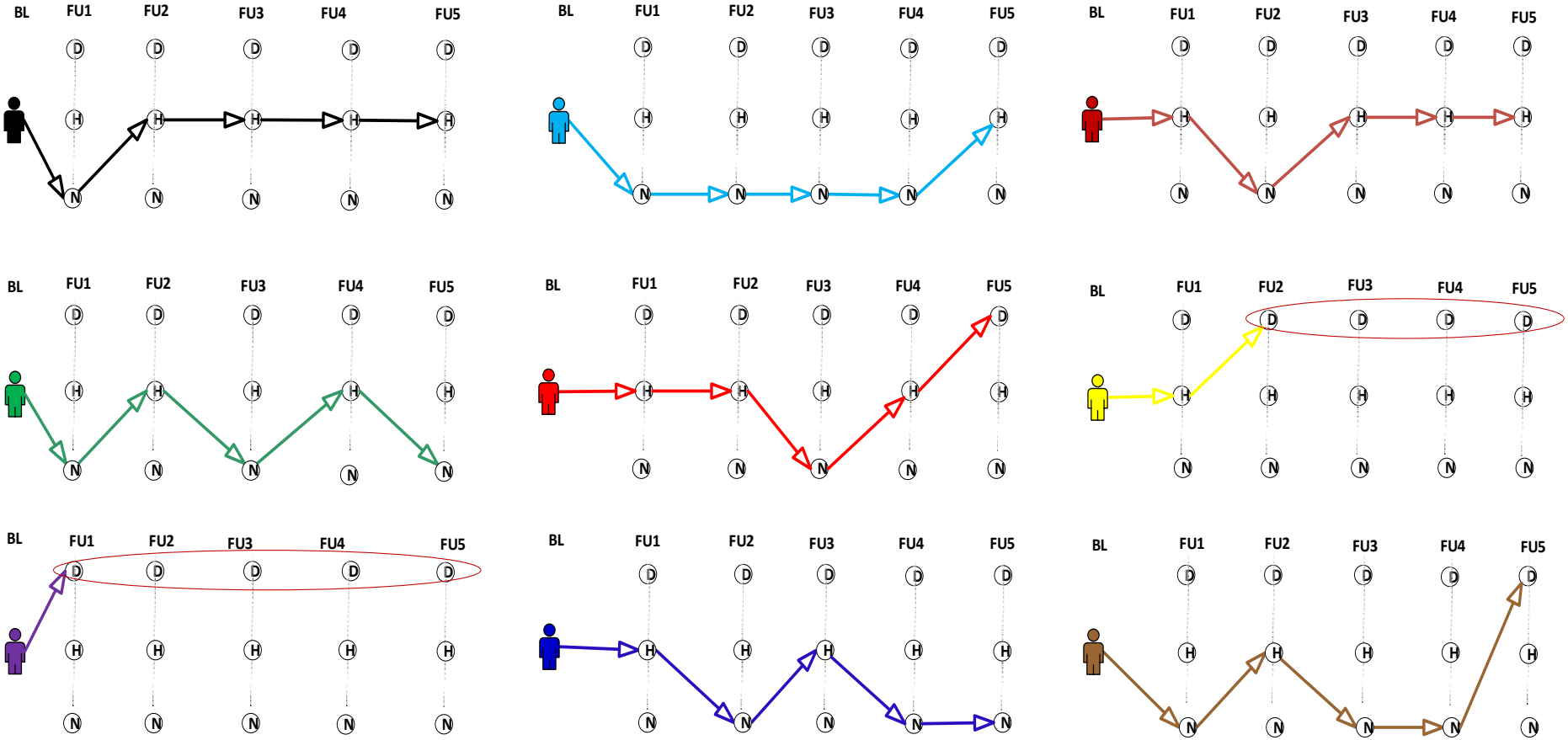


Figure 4.1: Shows graphically represents 9 cases of patient's healthcare journeyCapturing key interactions such as OPD follow-ups (O), admissions (H), neither (N) and Death event (D) over time. BL; baseline OPD, FU; follow up – interval (1, 2, 3, 4 and 5)

Figure 4.2 outlines a workflow for development of risk modelling. It outlines a sequential process beginning with 'Data exploration', where categorical and continuous variables are identified, moving onto 'Data handling', which includes steps for dealing with missing data through imputation. The workflow then progresses to 'Classification algorithms', featuring various methods such as Decision Trees, Bayesian (i.e., Naïve Bayes classifier), k-Nearest Neighbors (kNN), and Support Vector Machine (SVM). The final stage is development of risk prediction models and then 'Evaluation against', where static models and patient data models (PDM) are used for assessment.

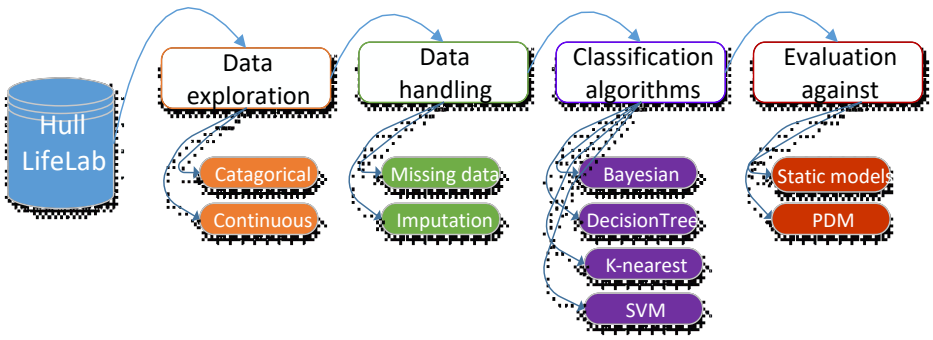


Figure 4.2: Data analysis workflow, from initial exploration to naïve Bayes classification & model development

4.2 **Hull LifeLab database**

The HLL database is organised into various clinical modules focusing on HF. These modules are interconnected via a central table called '***tpMain***'. Each patient is assigned a unique

identifier (*Link_ID*) and each clinical event is associated with a specific date (*Reference_Date*), which enable precise navigation and aggregation of individual patient data across various clinical modules. The events are recorded in a time-series format, meaning that each event is logged with a timestamp that allows for tracking the sequence and timing of events. This enables the analysis of how patients' conditions and treatments evolve, which is particularly important for longitudinal studies and developing predictive models for clinical outcomes. The described structure includes only the clinical modules that are relevant to the objectives of this study being discussed. The database is designed with robust security measures and access control mechanisms to safeguard patient confidentiality. These provisions ensure that sensitive patient information is protected, aligning with high privacy and data protection standards. The structure (Figure 4.3) is categorised in out-patient (OPD) clinical modules and risk events (hospital admissions and mortality data).

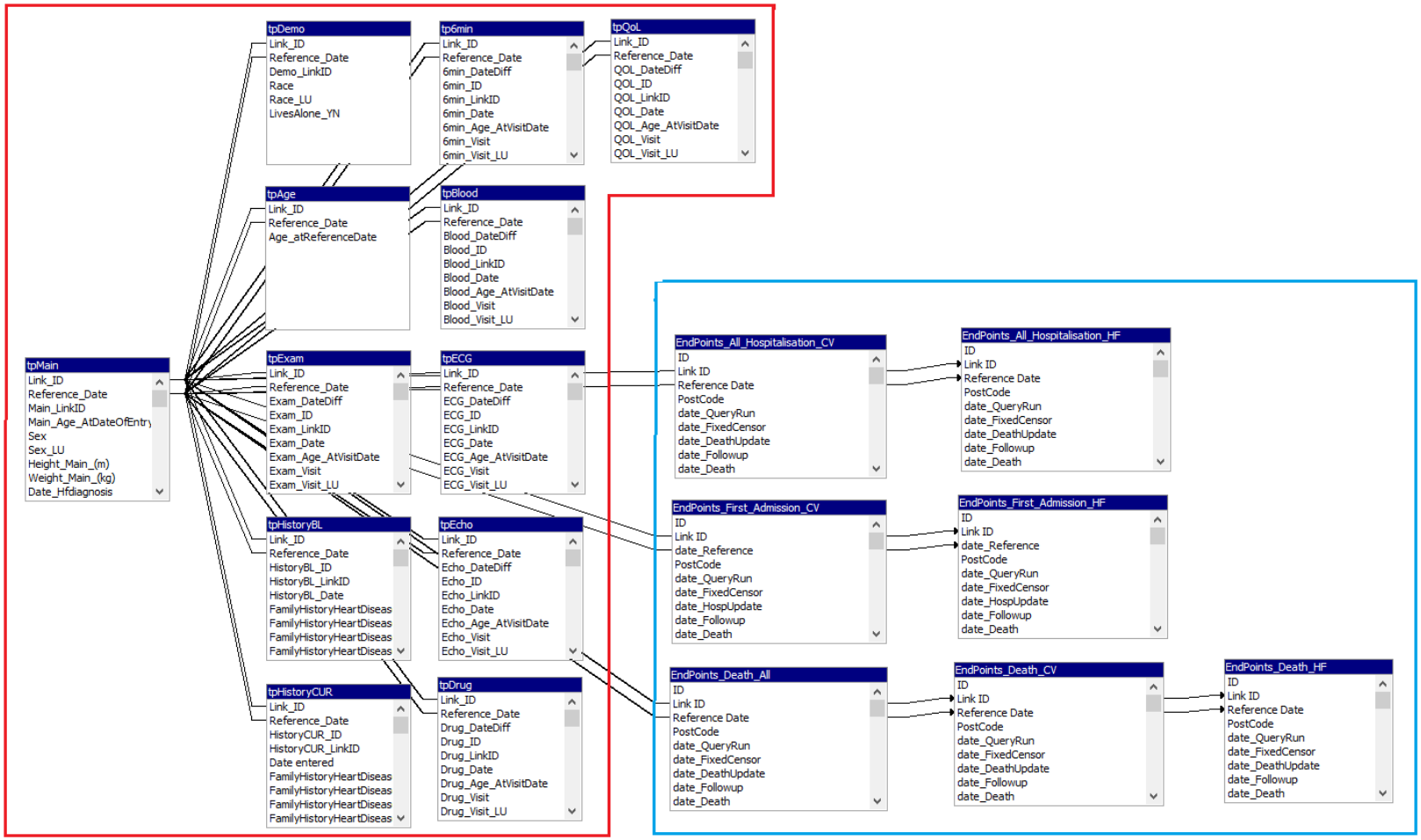


Figure 4.3: Structure of the HLL Clinical Modules for Analysis. The data presented in the HLL clinical modules is fully anonymised, ensuring patient privacy. Each individual in the dataset is uniquely identified by a "Link_ID". Together, the "Link_ID" and the "Reference_Date" correspond to unique entries within the dataset. Endpoints such as hospital admissions are linked with the patients' anonymized records by matching through an encrypted NHS/Hey Number. All events in the database are logged in a chronological time-series manner, capturing the temporal sequence of healthcare events and interventions for each patient.

Table 4.1 shows the table name, description and number of covariates of each clinical module. The number included the table joining keys.

Module Type	Table name	Table description	Number of variables
OPD tables	*tpMain*	Patient detail	28
	tpDemo	Socio-demographical	17
	tpAge	Age at each visit	3
	tpExam	Medical examination	57
	tpHistoryBL	Baseline medical history	66
	tpHistoryCUR	Current medical history	66
	tp6min	6 minutes' walk test	13
	tpBlood	Blood tests results	67
	tpECG	ECG readings	40
	tpEcho	Echo results	98
	tpDrug	Medication	138
	tpQoL	Quality of life questionnaire	165
Event data	All hospitalisation (end-point)	Hospital admission	42
	All cause deaths (end-point)	Mortality	90

Table 4.1: HLL tables, description of clinical module and number of covariates.

The number includes table-joining keys.

Between January 1st, 2000, and January 1st, 2017, 7,639 patients were referred to HLL services. Unless death occurs, a follow-up period of at least 24 months will be considered for all patients. The study will end on January 1st, 2019. The end of follow-up will be at the earliest of either the end of the study period or the mortality date.

Table 4.2 outlines the total patients and total records counts in each table. Variations in record counts across these tables indicate that not all modules are required for every patient at every visit. tpMain, tpAge, tpDemo, tpHistoryBL, and tpHistoryCUR contain a single record per patient. Each table contains a larger number of variables, proving a comprehensive dataset for developing risk models tailored to the research objective.

Module Type	Table name	Total population	Number of records
OPD modules	*tpMain*	7,639	7,639
	tpDemo	7,639	7,639
	tpAge	7,639	29,637
	tpExam	7,610	31,125
	tpHistoryBL	7,628	7,628
	tpHistoryCUR	7,628	7,628
	tp6min	3,311	8,962
	tpBlood	7,166	27,938
	tpECG	7,521	25,933
	tpEcho	7,570	20,198
	tpDrug	7,487	28,912
	tpQoL	6,024	19,262
Risk Event	All hospitalisation (end-point)	7,385	77,172
	All cause deaths (end-point)		

Table 4.2: Total population and number of records in each HLL table

Blood test results are extensive, with 27,938 records for 7,166 patients (Table 4.2). Regarding risk events, there are 77,172 hospital admissions recorded for 7,385 patients, suggesting that some patients were admitted multiple times. Considerable effort is required to align patient time series data across tables — a challenge that is tackled in a subsequent section.

Table 4.3 depicts the timing and total number of hospital admissions relative to the BL OPD visit. It categorizes admissions into three timing groups: 'Pre', 'Index', and 'Post', representing admissions before, on the day of, and after the BL OPD visit, respectively. Specifically, there were 32,286 'Pre' admissions before the BL visit, 543 'Index' admissions on the day of the visit, and 44,346 'Post' admissions following the BL visit.

Hospitalization Timing	**Total readmissions**
Pre	32,286
Index	543
Post	44,346
Total	**77,175**

Table 4.3: Total admissions of HLL population

'Pre' represents patients' admissions before the HLL baseline (BL) visit. 'Index' illustrates the day of the BL visit. 'Post' represents hospitalisation after the BL.

My study primarily focuses on post-baseline OPD cardiovascular (CV) readmissions. Table 4.4 outlines 44,346 (post) admissions, differentiating CV from non-CV caused of readmissions. Of the 6,732 readmitted patients, 25,699 admissions were cardiovascular (CV) events. These were subdivided into 7,667 heart failure (HF) and 118,032 other CV admissions. Non-CV causes accounted for 18,647 admissions, indicating that some patients experienced multiple admissions within the specified time window.

Admission type (post- BL OPD)	**Total number of admissions**	
Cardiovascular	25,699	
- HF Related		7,667
- Other CV		18,032
Non-Cardiovascular	18,647	

Table 4.4: Breakdown of post admissions

Based on whether they are related to cardiovascular issues and non-cardiovascular. Cardiovascular admissions are further divided into HF related and other cardiovascular related issues

Table 4.5 displays the distribution of the patients who had post-CV admissions. Out of 5,948 patients with post-CV admissions, 774 had exclusively HF admissions, 2862 had other than HF admissions, and 2,312 had a mix of both HF and other-CV admissions. This data provides insights on the prevalence of hospital events, particularly highlighting the incidence of cardiovascular-related readmissions.

Patient with HF admission only	Patients with other CV only	Patients with Both HF and Other Admissions	Total Patients
774	2862	2312	5948

Table 4.5: Distribution of patients based on their (post) admissions

4.3 Clinical data and its challenges

Compiling comprehensive medical histories into time series data from real-world clinical systems poses significant challenges. The dispersion of data across various tables (as shown in Table 4.2), for example, ECG, Echo, and Blood tests with HES and ONS, complicates the linking process in HLL's data. It requires extensive validation to ensure accurate data linking, clear definitions, appropriate labelling, and precise coding practices. The complexity of these tasks increases with the growing volume of data over time.

OPD visits and admissions occur at varying intervals, resulting in data entries that do not follow a uniform timeline. This complicates the data management and integration. The lack of uniformity is further complicated due to varying types of events such as cardiovascular (CV) versus non-CV admissions. If these issues are not effectively managed, they can undermine the reliability and validity of research outcomes. Understanding this, HLL has established systematic methods and protocols across all phases of data handling to ensure effective management of these complexities. This enhances the accuracy and reliability of research outcomes. Please refer Figure 4.1 for better understanding.

4.3.1 Misaligned sample data

Time series health-care data are considered misaligned when different types of clinical information were not recorded at the same time points or intervals. This misalignment can occur in two ways:

- In the first scenario (Figure 4.4), there is intra-individual misalignment when clinical data from different tables (like ECG readings, echo results, and blood tests) for the same patient are not collected at the same time points. For instance, ECG readings might be taken at regular intervals, whereas echo results and blood tests might not, resulting in uneven data distribution over the patient's timeline. Additionally, not all clinical information must be collected or completed during each patient visit, which can lead to incomplete datasets for certain time points.
- In the second scenario (Figure 4.5), inter-individual misalignment is illustrated, where clinical assessment timing is inconsistent across different patients within the research population. This indicating a lack of uniformity in scheduling and conducting follow-up events. For example, while one patient might have a follow-up after 4 months, another might only have theirs after a year. A practical demonstration of misalignment can be given during my viva.

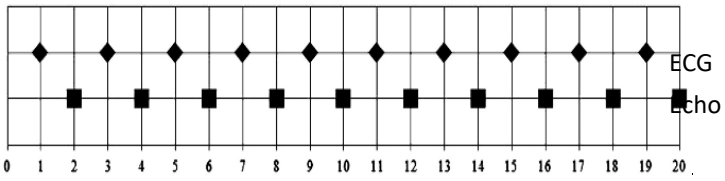


Figure 4.4: Misaligned sampled data – Time orientated

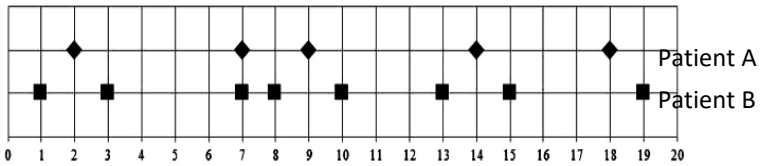


Figure 4.5: Misaligned sampled data – Follow up events

HLL exhibits similar misaligned patterns as described above. To address these challenges specialised data management tools were developed to align, and integrate information from different tables and time points.

A visit audit tool, as illustrated in Figure 4.6, is designed to tackle the issue of misaligned patient level data (intra-individual) from various OPD visits. This tool aligns clinic records and assigns a visit numbers based on criteria set by heart failure experts to address different research questions. As an example, the figure shows the first two years of patient's (link_ID: 68) clinic visits, displaying each visit date in the left column and the number of months since their first visit beside.

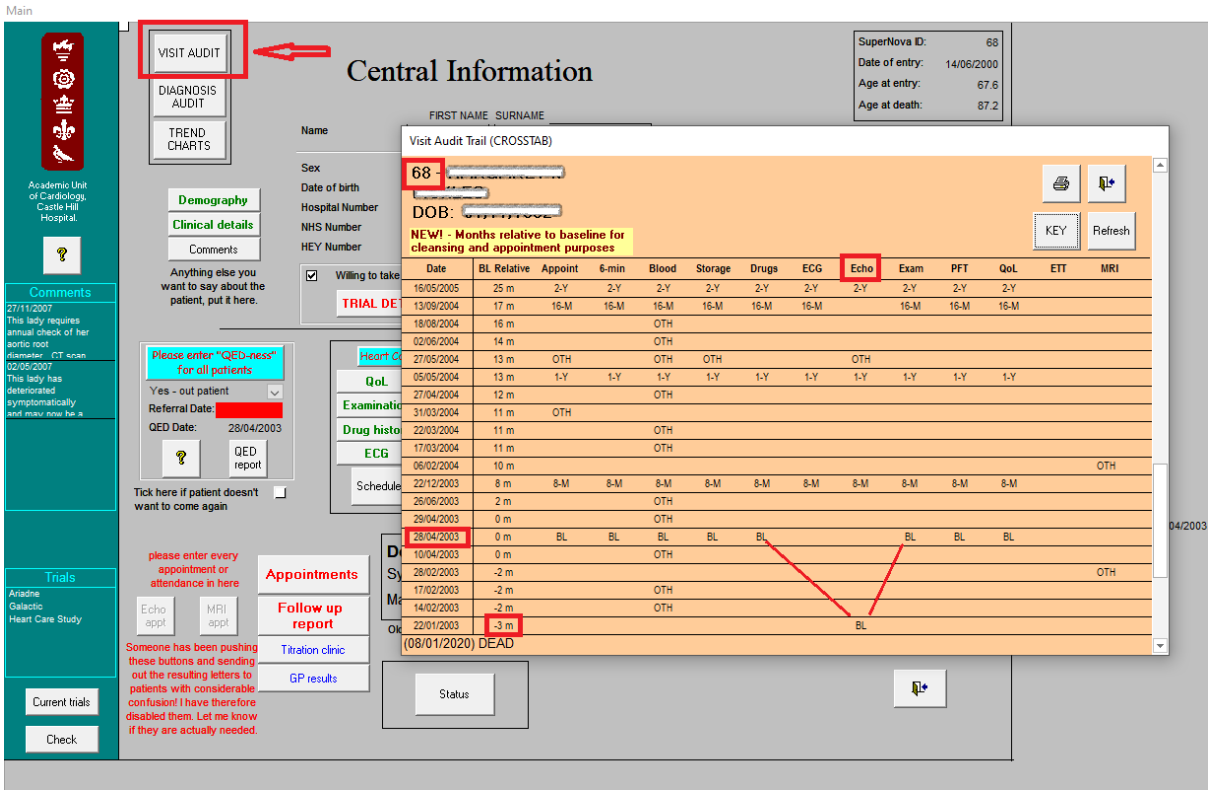


Figure 4.6: Visit Audit tool. The tool is designed to address the first scenario (misaligned sample data – time oriented), when patient (intra-individual) level data from different OPD visits don't line up properly.

This tool is designed to number each clinic visit according to predefined rules. If a module is missing for a specific date, the tool searches for the nearest date with available data and assigns the visit number accordingly. For instance, the figure demonstrates that an echo occurring within ±4 months of the baseline visit is labelled as the BL echo. This method streamlines the process by creating a series of visits, each represented in a single row.

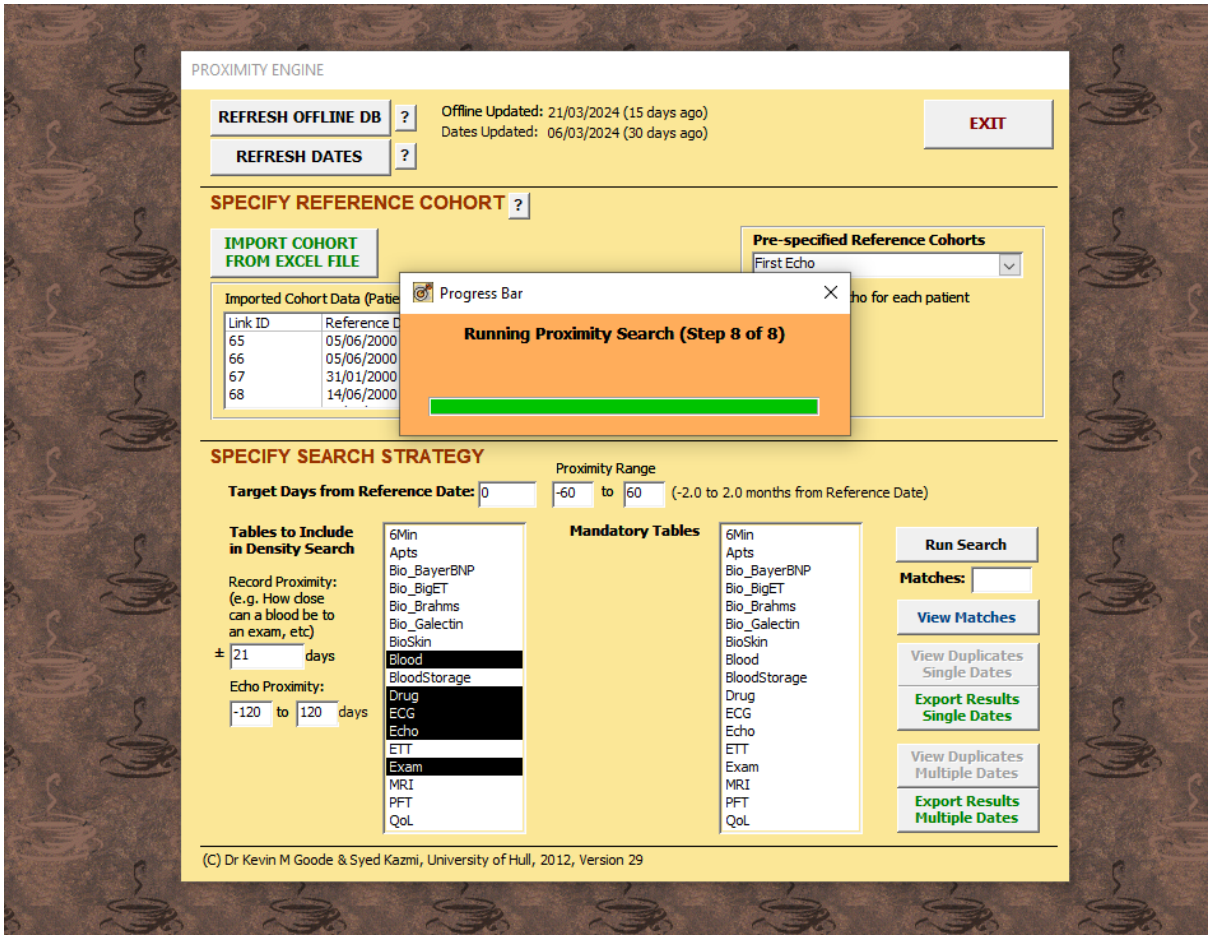


Figure 4.7: Proximity Engine

The tool select a study cohort based on criteria aligned with the research goals. It then conducts a targeted search within the modules, using time windows tailored to study requirement.

To address inter-individual misalignment in research populations three tools streamline the dataset for analysis. First, a visit audit tool organizes all OPD visits and associated data for each patient in chronological order as closely as possible. Next, the proximity engine (Figure 4.7) is used to select various patient cohort based on criteria aligned with the research objective. For instance, the figure demonstrates how the tool identifies patients' IDs (link_ID) and dates of Echo (Echo_Dates). It then cross-reference other HLL tables and searching for dates within

specific ranges (e.g., ± 60 days for ecg, medication, examination; ± 21 days for blood; ± 120 for echo). This crucial step consolidates all visit dates from various tables into a unified record per line, illustrated by Figure 4.6. IDs (*Link_ID*) and dates (*Reference_Date*). Echo_Dates are considered as *Reference_Date* in this example.

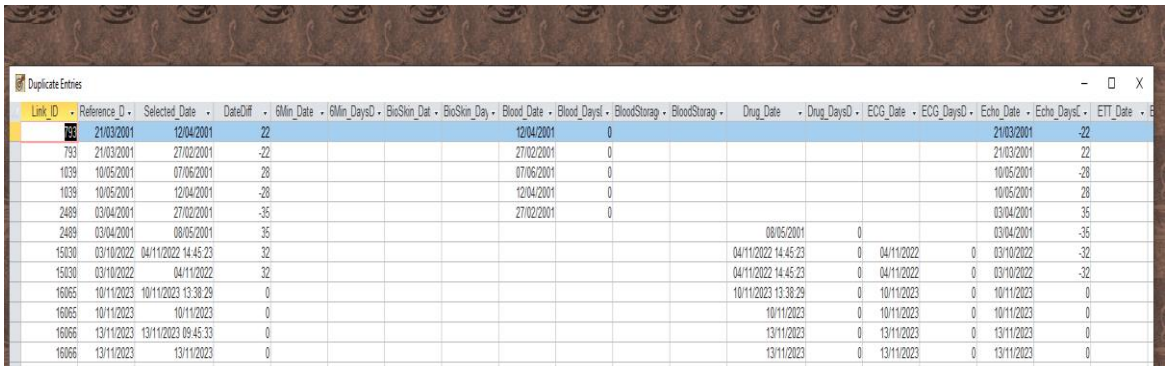


Figure 4.8: Duplicate records- obtained through the proximity of Data Engine.

The list is then processed through the Data Engine (Figure 4.9), which retrieves data from all selected tables, resulting in a comprehensive list of the study population's data. The data is organised to display each patient's records for every visit in a single row.

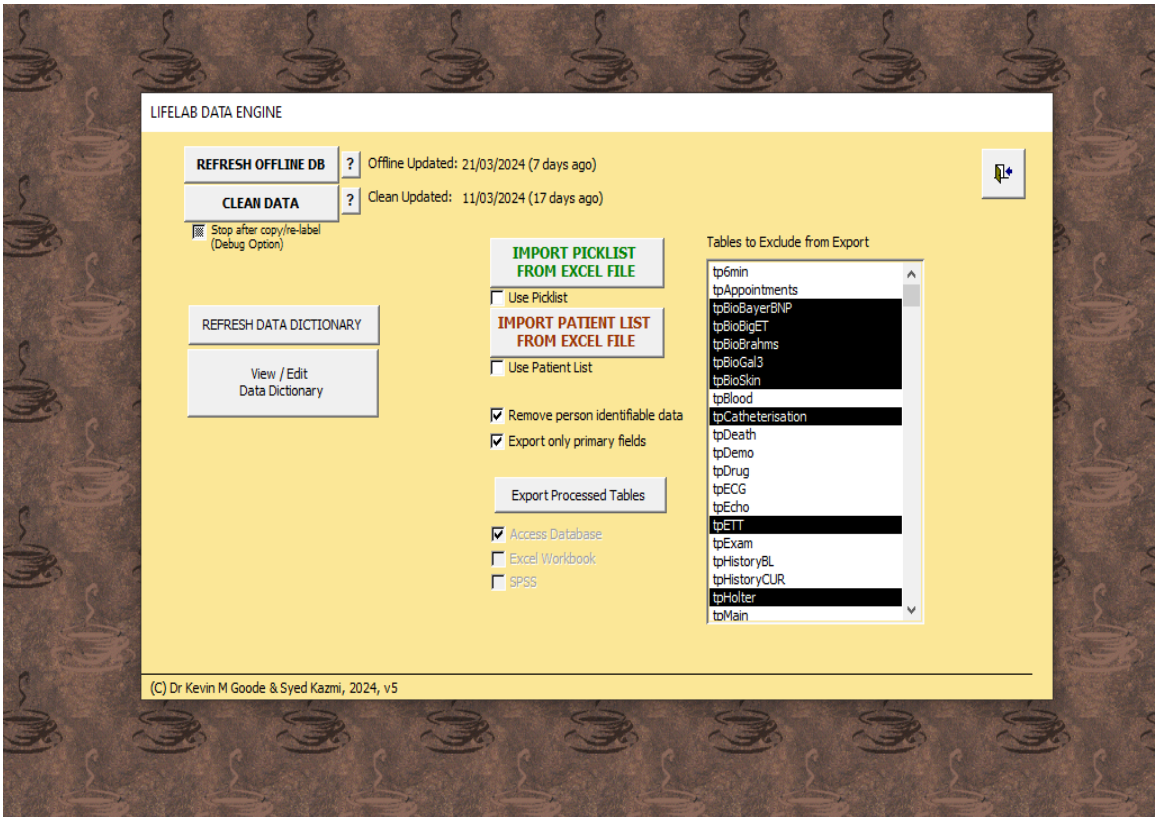


Figure 4.9: Data Engine. The tool is designed to pulls together all necessary clinical variables from the respective modules for the selected cohort obtained through proximity engine.

The final step involve gathering events data (e.g. for end-points analysis – all CV admissions and all-cause mortality) of selected cohort. This is achieved through End-Point Engine (depicted in Figure 4.10).

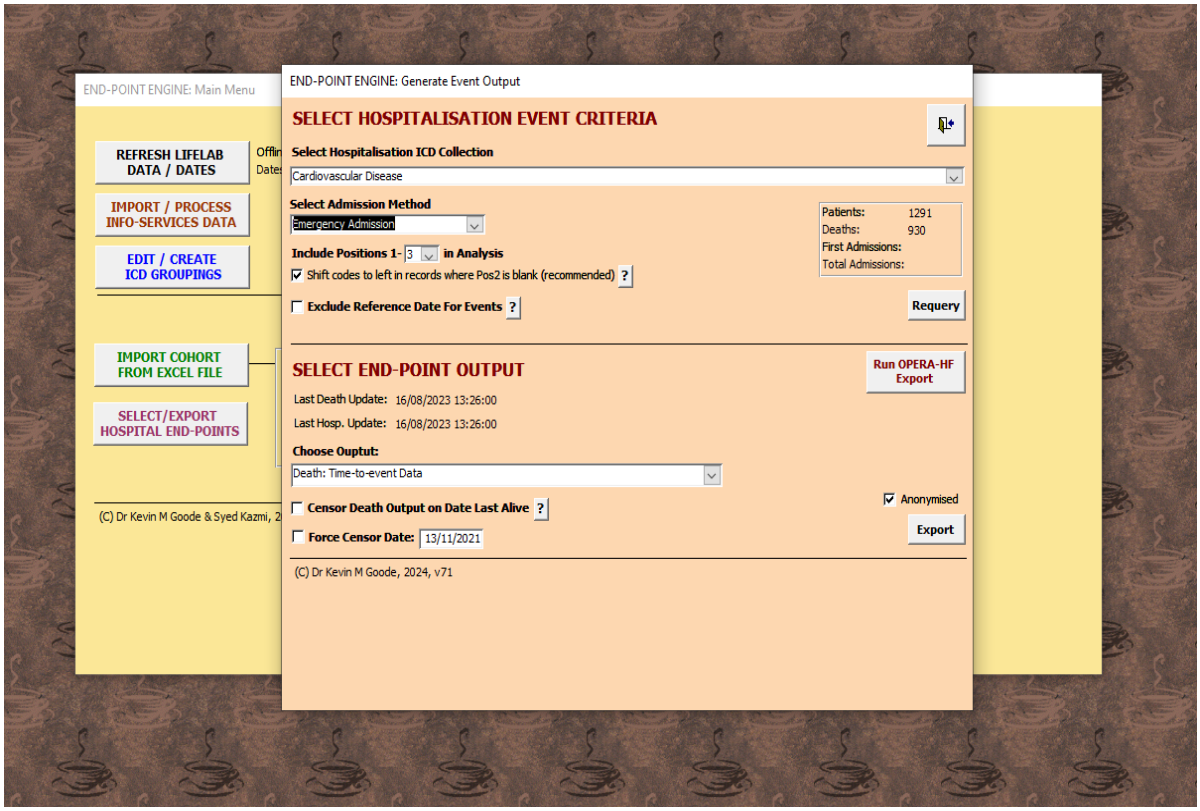


Figure 4.10: End-Point Engine. Tool designed to extracts events (end-points – all hospital admissions and mortality).

With complete data including all OPD visits, hospital readmissions, and mortality information, researchers can analyse HF progression over time. They explore patient health trajectories, observe if conditions worsen, improve, or remain stable, and identify factors influencing these outcomes.

4.3.2 Repetitive records and missing values

The data from the final three columns of Table 4.2 reveal two key insights:

- Not every clinical test (e.g., *tpBlood* or *tpEcho*) are performed at every OPD visit for each and;
- Within each table there's a risk of repetitive records if not carefully aligned, such as one record containing maximum information and another with only one clinical variable noted.

Records appearing as repetitive may not be true repetitive:

- Visits might have dates very close to one another, giving the impression of repetition when they are separate visits within a brief period.
- Sometimes within the same clinical table different variables may be recorded on different dates, resulting in entries close together that represent various aspects of one clinical information over multiple visits.
    - These closely time spaced records, reflecting partial completion of clinical tables across different dates, could mistakenly be considered repetitive without careful examination of the timing and context of each entry.

However, whether dealing with repetitive records or missing values, both scenarios substantially impact the accuracy and effectiveness of predictive risk modelling. This can impact the model's ability to forecast outcomes accurately (Cismondi et al., 2013).

The two most commonly employed strategies for handling missing data involve imputing (filling in) the missing values or deleting the records with missing information (Li Peng et al, 2015). Each strategy has its own set of pros and cons. The choice between these methods often depends on the specific circumstances of the study, but the guidelines for preference are not always clear (Brock et al., 2008). Imputation can cause bias, and deletion can cause bias and statistical power loss (Cismondi et al., 2013; Li Peng et al, 2015). When

using longitudinal data, the direct deletion method can eliminate all the other clinical variable collected at a particular time. This can impact the completeness of the patient's health record and potentially affect the accuracy of disease trend predictions (Li Peng et al, 2015).

The simplest methods for handling missing values are Mean or Median substitution, which replace the missing values by the mean or median of all the observational values. Other common imputation methods for longitudinal clinical data include: Last Observation Carried Forward (LOCF), Linear Interpolation, Predictive Mean Matching, Growth Curve Modelling and Multiple Imputation (Poolsawad et al., 2012; Zhang et al., 2012).

HLL has set up robust procedures and system, as outlined in section 4.3.1, to manage missing and repetitive records. In this study, I chose not to use imputation methods. Instead, with advice from HF disease experts, I applied specific criteria via the visit audit and proximity engines. For example, I filled gaps by identifying and using the nearest available record to the reference date. As shown in Table 4.2, out of the 19,019 echocardiogram (tpEcho) records, 669 were re-selected because they fell within a ±120 days window of the reference date. This method allowed us to use actual patient data to address gaps in the dataset.

4.3.3 Non-normal distribution

In clinical data analysis, the assumption of a symmetric bell-shaped (normal) distribution is rarely met. Clinical datasets frequently exhibit significant skewness, which can lead to erroneous conclusions of statistical analysis. Various techniques exist to check for normality. Before moving to detailed analyses, understanding fundamental concepts and recognizing different data shapes seen in initial exploration is important. This understanding is essential for selecting appropriate methods to address non-normality.

Observing the shape of the data distribution and plotting (visual representation) is a better way to describe uncertainties in the dataset (David, 2011; Anderson & Druker, 2013; Anderson, 2014). Data distribution is typically categorized into three main shapes: symmetric, left-skewed, and right-skewed, as shown below.

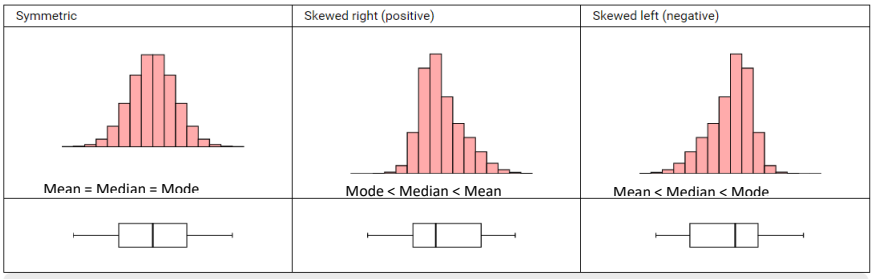


Figure 4.11: Distribution shapes and Skewness

Clinical variables assessing health status often show asymmetric, long-tailed distributions, skewing right (Counsell et al., 2011). This means that most individuals have low values, while only a few have large values, often indicating a specific illness. To analyse non-normal distributions, transforming the data might be needed. Transformations like logarithmic, square root, Box-Cox, and reciprocal help make skewed data resemble a normal distribution. Statistical analyses often use distributions like Binomial, Poisson, Uniform, Normal

(Gaussian), and Negative Exponential (Glasgow Caledonian University, 2012). The Normal distribution, vital for continuous variables, shows data trends through its bell-shaped curve, which extends infinitely in both directions, as shown in Figure 4.11.

As illustrated in Figure 4.12, the curve of a normal distribution approaches but never touches the horizontal axis, extending beyond about 3 standard deviations from the mean. The probability density function's equation (4.1) involves parameters μ (mean) and σ (standard deviation), where approximately 68% of values lie within one standard deviation (SD) of the mean. The SD measures the average distance of values from the mean, squared and then square-rooted.

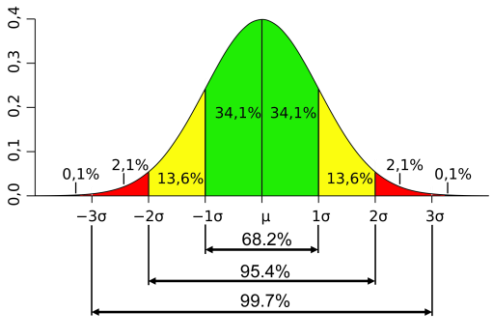


Figure 4.12: Standard normal distribution.

We meet standard normal distributions (SDs) later in the thesis represented as Z-scores; chapter 8.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty;\ -\infty < \mu < \infty, \sigma > 0 \tag{4.1}$$

Equation 4.1: Probability density function's equation

68% of all values will lie within $\mu \pm \sigma$
95% of all values will lie within $\mu \pm 2\sigma$
99% of all values will lie within $\mu \pm 3\sigma$

When data doesn't follow a normal distribution, two steps are necessary: i) identify the cause of the non-normality, and ii) dealing it by transforming the data to normalize it or using non-parametric statistical methods. Causes of non-normality include:

- Outliers, which are extreme data points that can distort the distribution.
- Overlap of two or more processes, where data from different sources or factors mix, often seen in medical settings with inputs from multiple sources.
- Data values near zero or a natural limit, leading to skewed distributions due to clustering at the lower or upper bounds.

4.4 Study dataset

I conducted three distinct analyses in the study, focusing on patients' state-to-state transitions. Each using datasets tailored to transition state diagram showing different trajectories to address specific research questions. I organised and streamlined patient states to display disease progression (detailed in later chapters). Each approach is visually represented through individual patient case vignettes, highlighting state transitions of each model.

Given the vast amount of data in the HLL database, focusing on relevant clinical variables can reduce missing data and incomplete records. Careful selection of variables is essential for dynamic risk modelling. I presented baseline characteristics and demographics for the dataset, offering insights into the patient population and setting the stage for deeper

analyses. Data on events were collected for an additional two to five years beyond the initial study period, enabling verification of the model outcome.

- **Selection of patients**

A cohort of 7,639 patients, referred to the OPD clinic between January 2000 and January 2017 was enrolled. Each patient had a baseline visit documented, with subsequent follow-up visits, admissions, and mortality data recorded if applicable. The specific timeframes and the characteristics of the patient groups for each analysis are explained in the corresponding chapters of the analysis.

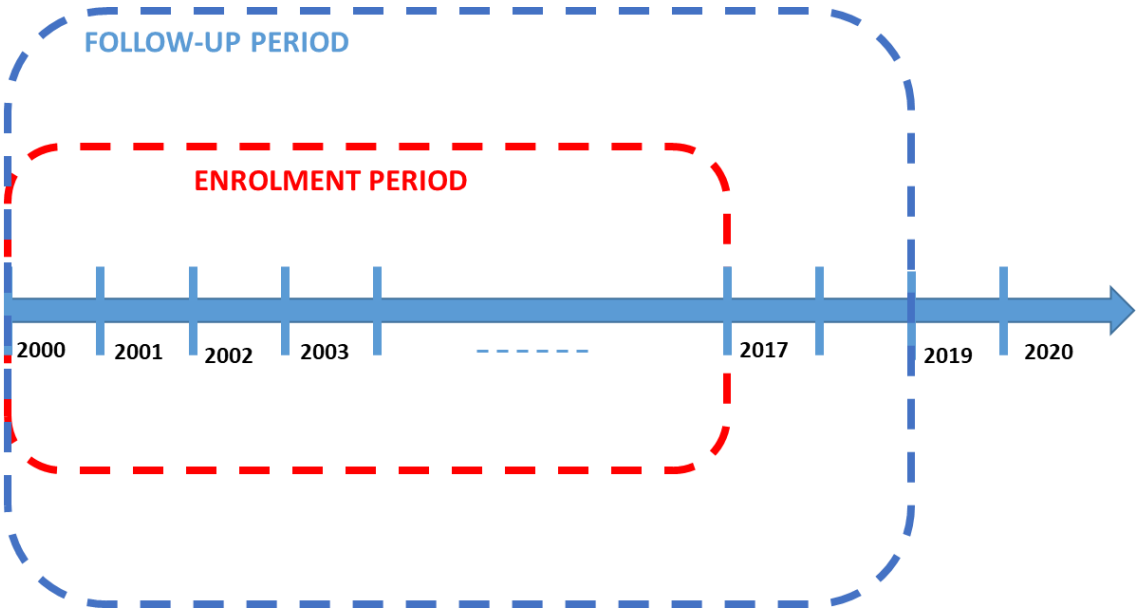


Figure 4.13: Enrolment and follow-up period

- **Covariate selection**

HLL contains a mixture of continuous and categorical data (nominal and ordinal). Explanatory covariates were selected for each analysis, which was chosen with input of heart failure disease experts and also based on research papers that were associated with outcome. (Bohacik et al., 2013; Pocock et al., 2013; Nikolaidou et al., 2018; Sokoreli et al., 2018; Zhang et al., 2018a; Koulaouzidis et al., 2019).

Covariates selected are: age, gender, body mass index (calculated as the weight in kilograms divided by height in meters squared (BMI), New York Heart Association (NYHA) class, systolic blood pressure (SBP), oedema, Left ventricular systolic dysfunction (Left ventricular (LV) impairment, EF Simpsons), electrocardiogram (heart rate (bpm), heart rhythm (sinus), Auricle diastole Auricular repolarization Ventricular depolarization Cardiac cycle (QRS)) blood tests (haemoglobin, sodium, potassium, urea, creatinine, albumin, , N-terminal pro–brain natriuretic peptide (NTproBNP)), HF medication (loop diuretic, furosemide equivalent daily doses, thiazide, beta-blocker (Bb), angiotensin-converting enzyme inhibitor (ACE), angiotensin receptor blocker (ARB), mineralocorticoid antagonist (MRA)).

4.4.1 Baseline demographic and clinical characteristics

Baseline (BL) means the first visit to out-patient (OPD) clinical from the time-based series data. A total of 7,639 patients referred to the clinic between January 2001 and January 2017 were enrolled in the study. 4,131 individuals attended follow-up (FU) out-patient clinics, while 3,508 have only BL visits. This indicates that approximately 54% of the participants returned for follow-up at some point, whereas the remaining 46% never did (Figure 4.14).

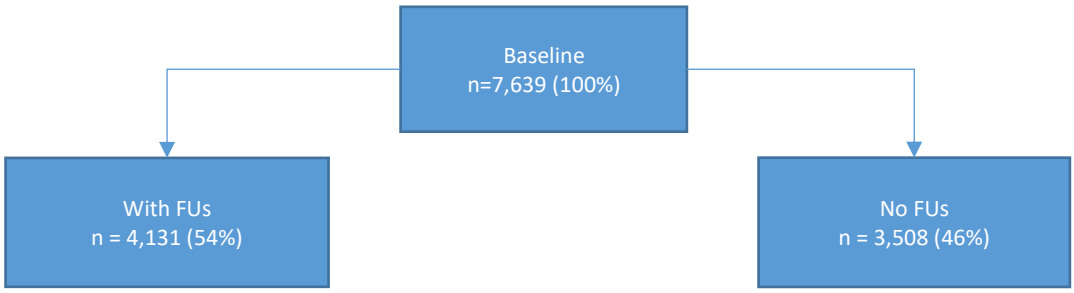


Figure 4.14: Study population

Table 4.6 provides patients' baseline characteristics, highlighting the differences between patients with OPD follow-up visits and those without OPD follow-up. This comparison aims to identify any unique baseline traits between the two groups.

Overall, people with one or more event had proportionately higher levels of oedema, NYHA class> III, worsening LVI, a greater proportion male. There were higher levels of NT-proBNP in those with at least one event. There were corresponding increases in ACE/ARBs, loop diuretics and beta-blockers. Age, heart rate, and sodium and potassium levels were similar in both groups. Both EH and SBP were lower.

	All baseline		With follow-up		No follow-up	
Variable	**Missing (n (%))**	**Total**	**Missing (n (%))**	**Total**	**Missing (n (%))**	**Total**
Total number of patients N (%)		7639 (100)		4131 (100)		3508 (100)
Demographics						
Age (years)		73.5 [66.6, 79.9]		73.4 [65.5, 79.8]		73.6 [65.7, 80.0]
≥75		3386 (44)		1815 (44)		1571 (45)
Men (%)		4492 (59)		2698 (65)		1794 (51)
BMI (kg/m2)	61	28.1 [24.7, 32.1]		28.2 [24.8, 32.2]	61	27.9 [24.7, 32]
NYHA class (%)	189				189	
I-II		5315 (70)		2784 (67)		2532 (72)
III-IV		2135 (28)		1347 (33)		789 (22)
Oedema (%)	681		311		370	
None		4087 (54)		2021 (49)		2066 (59)
Trace		944 (12)		600 (15)		346 (10)
Ankles		1398 (18)		927 (22)		471 (13)
Above ankles		529 (7)		272 (7)		255 (7)

Systolic BP (mmHg)	19	139.0 [122, 158]	4	135 [119, 154]	15	144 [127, 162]
Left ventricular systolic dysfunction						
LV Impairment	143		47		97	
None		3327 (44)		905 (22)		2422 (69)
Trivial		574 (8)		279 (7)		295 (8)
Mild		975 (13)		720 (17)		255 (7)
Worse		2620 (43.7)		2180 (53)		439 (13)
EF Simpson (%)	3506	47 [35, 58]	1764	39 [30, 50]	1742	56 [49, 62]
Findings on electrocardiogram						
Heart rate (bpm)	204	72 [62, 84]	47	72 [62, 85]	157	72 [60, 83]
Heart rhythm (Sinus %)		4976 (65)		2461 (60)		2515 (72)
QRS (ms)	351	98 [86, 118.0]	99	104 [90, 130]	252	92 [84, 106]
Blood test						
Haemoglobin (g/dL)	1370	13.4 [12.2, 14.5]	436	12.2 [11.2, 14.5]	934	12.2 [11.0, 13.3]
Sodium (mmol/L)	1224	139.0 [137, 140.0]	390	139.0 [137.0, 140.0]	834	139.0 [136.0, 140.0]
Potassium (mmol/L)	1265	4.3 [4.0, 4.6]	406	4.4 [4.1, 4.7]	859	4.3 [4.0, 4.6]
Urea (mmol/L)	1224	6.5 [5.0, 8.9]	390	6.9 [5.2, 9.3]	834	6.1 [4.7, 8.3]
Creatinine (umol/L	1232	97.0 [80.0, 121.0]	396	101.0 [83.0, 127.0]	836	91.0 [75.0, 115.0]
Albumin (g/L)	1326	38.0 [35.0, 40.0]	452	38.0 [36.0, 40.0]	874	38.0 [35.0, 40.0]
NT-proBNP (ng/L)	1166	815 [234, 2195]	271	1149 [468, 2689]	895	321 [315, 1336]
Heart failure medication						
Beta-blocker (%)		4333 (57)		3023 (73)		1310 (37)

Loop diuretic (%)	4858 (66)	3050 (74)	1518 (44)
ACE/ARB (%)	5166 (68)	3504 (85)	1662 (47)

Table 4.6: Baseline demographic and clinical characteristics of (N=7639) patients.

a) patients with follow-up (N = 4131)

b) patients with no follow-up (N = 3508).

*Continuous variables are presented as median (interquartile range), whereas categorical variables are expressed as numbers and percentage.

Abbreviations: CHF, Chronic heart failure; BMI, body mass index (calculated as the weight in kilograms divided by height in meters squared); NYHA, New York Heart Association. BP, blood pressure; LV, left ventricular; EF, ejection fraction; LVSD, left ventricular systolic dysfunction; NT-proBNP, N-terminal pro–brain natriuretic peptide; ACE, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; MRA, mineralocorticoid antagonist; ECG, electrocardiography; eGFR, estimated glomerular filtration rate; HF, heart failure; IQR, interquartile range; N, total number; %, percentage.

*NT-proBNP only became a clinical service during the data collection.

- All cause mortality

Table 4.7 presents the cumulative mortality rates for a cohort of 7,639 patients over various time intervals up to 60 months. The patients are categorised into two groups: those with follow-ups (FUs) and those without follow-ups (No FUs). At 30 days, the total mortality rate was 2% (177 patients), with only 4 (less than 1%) of these deaths occurring in the group with follow-ups and 173 (5%) in the group without follow-ups. Mortality is all-cause, otherwise specified.

All-cause deaths	**Total patients (n=7639)**	**Patient with FUs (n=4131)**	**Patient with No FUs (n=3508)**
30 days	177 (2%)	4 (0%)	173 (5%)
12 months	894 (12%)	283 (7%)	611 (17%)
24 months	1414 (19%)	612 (15%)	802 (23)
36 months	1852 (24%)	899 (22%)	953 (27)
48 months	2240 (29 %)	1144 (28%)	1096 (31)
60 months	2555 (33%)	1341 (32%)	1214 (35)

Table 4.7: Cumulative mortality rates for a cohort of 7,639 patients over various time intervals up to 60 months. The patients are categorized into two groups: those who had follow-ups (FUs) and those with no follow-ups (No FUs).

4.4.2 Data distribution

The dataset includes many continuous and categorical variables. Figure 4.15 illustrates the distribution of values of variables show in table 4.6. Majority of variables are right skewed as discussed section 4.3.3. For example, NT-proBNP is significantly skewed. A widely used method for handing skewed continuous data is logarithmic translation (Bland & Altman, 1996). This is further supported by (Olivier et al., 2008) which states that logarithmic transformation enables standard statistical methods to be applied to skewed data. The transformation is commonly uses the natural Log *e* or log10.

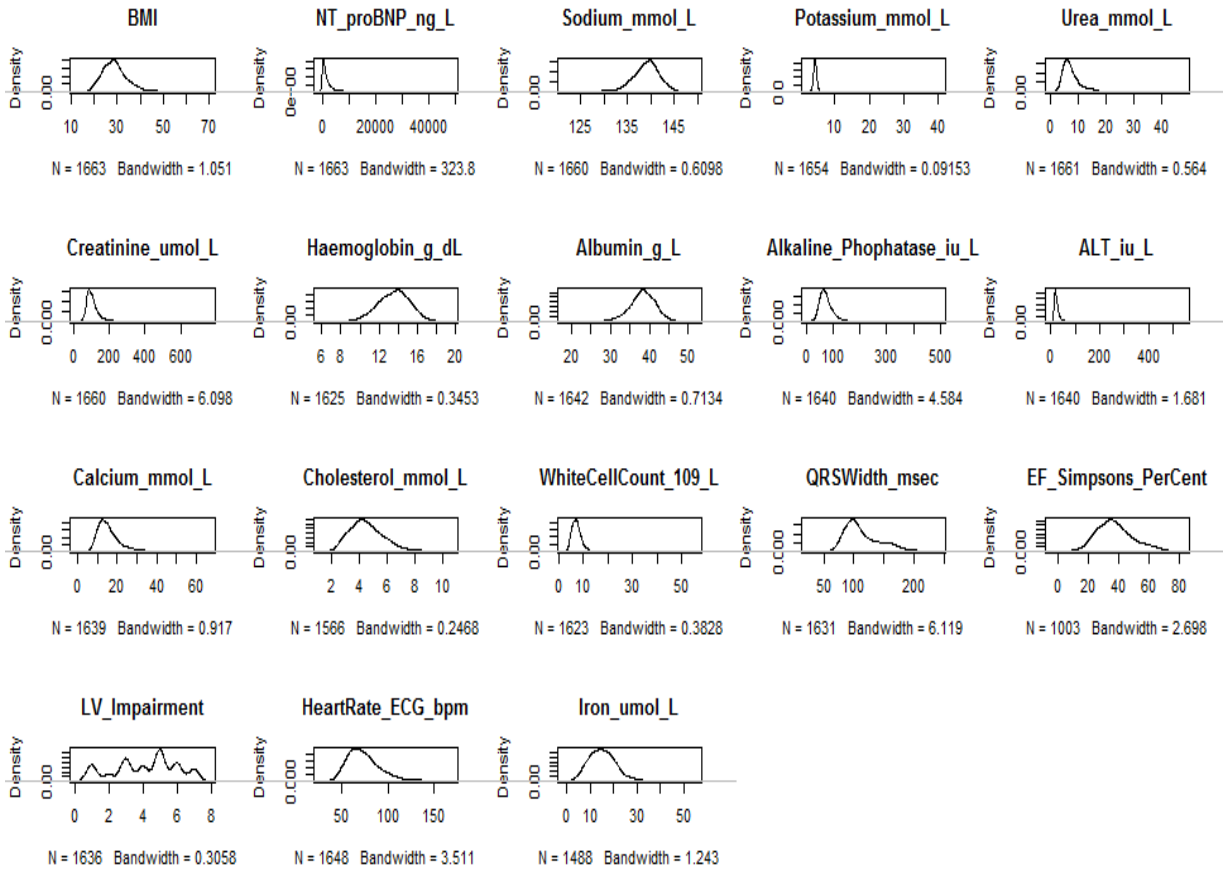


Figure 4.15: Distribution of baseline characteristics

4.4.3 Organisation of data

Within the framework of dynamic risk modelling, the study aims to track how patients' health evolves as they experience various risk phases—acute, unstable, and stable. Integrating OPDs, post CV admissions and mortality data introduces an added layer of complexity to the data analysis process. Which require sophisticated analytical techniques to interpret the data and uncover patterns, assess risk factors, and predict the probability of subsequent events. In this section, I initiate the organisation of the dataset into a structured format that sets the stage for dynamic risk analysis in later chapters.

- OPDs, admission (hospitalisation) and mortality

Table 4.8 summarizes the distribution of data up-to 24 months. Patient can have more than one event per interval. Table 4.9 refines the dataset by selecting one event within 4 monthly intervals.

No	Interval	4 monthly window after BL	OPD		Admissions (post - CV)		Death
			Total FUs	Total patients	Total FUs	Total patients	
1	1M	<= 2	757	752	1050	900	177
2	4M	>2 to <=6	2983	2970	1747	1326	268
3	8M	>6 to <=10	1175	1168	1427	1124	235
4	12M	>10 to <=14	2330	2320	1313	1005	214
5	16M	>14 to <=18	1408	1400	1126	877	166
6	20M	>18 to <=22	723	719	1132	887	195
7	24M	>22 to <=26	1772	1763	1024	808	159
		Total	**11148**	**11092**	**8819**	**6927**	**1414**

Table 4.8: Total OPD FUs, admissions (post - CV) and death

Based on intervals adjusted for time windows after the BL OPD visit. Table description is as follow: The first column (No) shows sequence of interval (second column). The third columns (4 monthly window) shows monthly window after BL to assign the interval. E.g.,: second row show that within (>2 to <=6) widow total 2983 OPD FUs for 2970 patients, 1747 admissions for 1326 patients and 268 deaths were observed. The time intervals are used in later sections and chapters.

No	Interval	4 monthly window after BL	OPD		Admissions (post - CV)		Death
			Total FUs	Total patients	Total FUs	Total patients	
1	1M	<= 2	752	752	860	860	177
2	4M	>2 to <=6	2970	2970	1234	1234	268
3	8M	>6 to <=10	1168	1168	1032	1032	235
4	12M	>10 to <=14	2320	2320	938	938	214
5	16M	>14 to <=18	1400	1400	816	816	166
6	20M	>18 to <=22	719	719	836	836	195
7	24M	>22 to <=26	1763	1763	738	738	159
		Total	**11092**	**11092**	**6454**	**6454**	**1414**

Table 4.9: 1 event (OPD, admission and death) per 4 monthly intervalSee Table 4.8 for a table description.

4.5 Conclusion

This chapter has introduced the HLL. It is outlining its structure and the clinical information it provides for analysis. I have discussed the challenges of managing clinical data and offered strategies and examples to address these issues. It is including visually represented through individual patient case vignettes, highlighting state transitions, showing the case studies to clarify my methods. The chapter aims to develop basis for new predictive risk models that accurately reflect patient outcomes and predict future events in those with CHF, employing a dynamic approach. This involves a data-driven method aligned with a risk trajectory framework, categorizing health states as acute, unstable, or stable. I can track the patient's risk trajectory over time by assigning health risk states —like OPD visits, hospitalizations, and mortality. In conclusion, the research within this chapter sets the stage for a nuanced exploration of chronic heart failure, providing the groundwork for innovative risk assessment and predictive modelling that could significantly impact patient care.

Chapter 5 Naïve modelling process

As discussed in the previous chapters the HF illness trajectory is complex and dependent on several crucial aspects that need to be understood. For example, uncovering the underlying progression pattern, predicting the total number of events in a particular state before transitioning to another state, and estimating the duration of individuals' lives (length of life) for those with HF. Machine learning techniques have received considerable attention for making predictions with regards to time to event especially when clinical datasets contained repeated observations as highlighted in (Zupan et al., 2000; Wolfson et al., 2015). For deeper exploration of how these methods have been especially within my context, see chapters 2 and 3.

This chapter aims to predict the first event not the subsequent events and to investigate the traditional machine learning techniques to analyse disease trajectories in patients with CHF. For this purpose I used classification methods to map the progression of patient conditions over time. As will be discussed in the later sections of this chapter, these methods alone are often insufficient to capture a complete picture of HF disease progression. Clinical decision-making is not a one-time event, it is an ongoing process that may require adjustments based on evolving patient conditions, new information, or changes in treatment guidelines. It is important to know that how patient's health state changes or the duration of different disease phases (acute, unstable and stable). Such insights are important for determining each patient's most appropriate course of action. Such adjustments are necessary to accommodate as these directly impact patient outcomes, quality of life and even survival.

5.1 Classification

Data mining is a machine learning (ML) process that involve extracting and uncovering valuable knowledge from vast resources of data. The main components for data mining are: Classification, Sequence Analysis, Clustering and Association rule learning (Kesavaraj & Sukumaran, 2013). Classification technique that is used to identify the risk factors that lead to classifying each record into a predefined set of classes and groups (Kesavaraj & Sukumaran, 2013). These models use learned patterned from the training data to assign instances to the most appropriate class.

Many studies (Kheirbek et al., 2013; Comstock Barker & Scherer, 2017; Maheswari & Pitchai, 2019; Gupta et al., 2020; Sai Krishna Reddy et al., 2022) on HF rely on classification methods like Naïve Bayes and decision trees. Some of them are considered prognostic models or diagnostic models. Figure 5.1 is taken from the study (Ohu et al., 2020) to represents a broad classification types of learning, methods utilised in ML and category of algorithm.

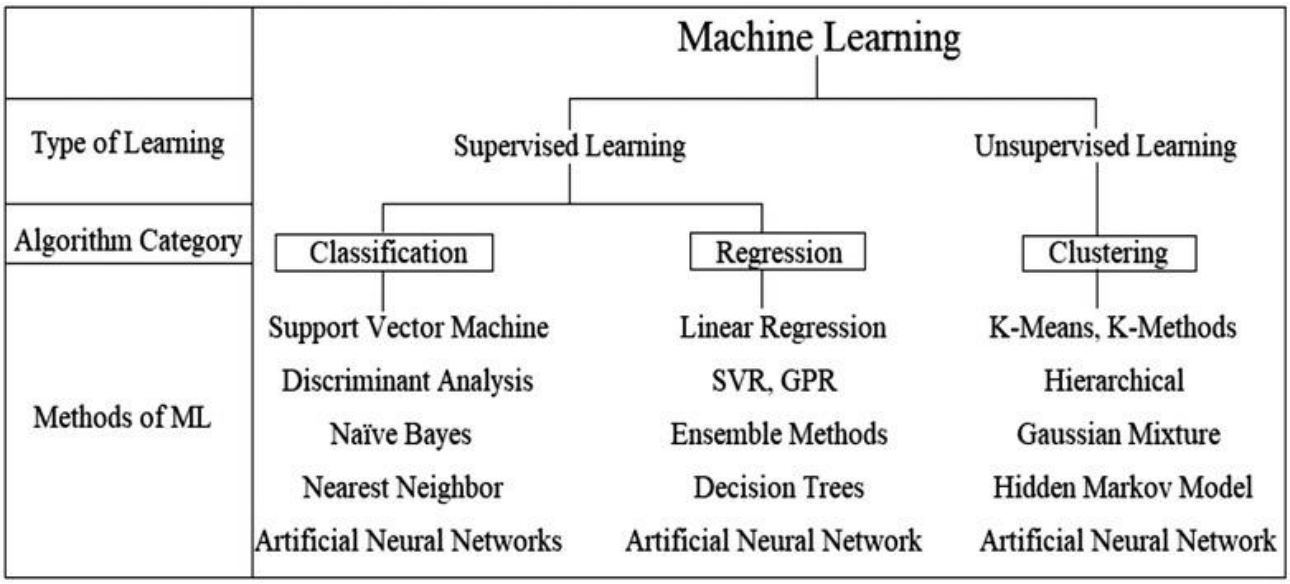


Figure 5.1: Categorisation of ML models.SVR (support vector regression); GPR (Gaussian process regression) (Ohu et al., 2020)

The principle of ML techniques is to train a model on different subsets of data. This process involves dividing the labelled dataset into two main subsets: the training and the test datasets. Classification in ML divided into binary, multi-class, multi-labelled and hierarchical tasks. Naïve Bayes, Bayes Net and RBF Network classifier are applied to the HLL dataset to classify relationships into two different predefined risks: a) Alive or dead (binary value) and, b) out-patient visit or hospitalisation or dead. Classification methods like Naïve Bayes rely on conditional independence assumptions that the presence of a specific feature is assumed to be independent of the presence of other features, given the class label being predicted. Similar kind of assumptions can be made with Markov Chains as can be seen in Chapter 6. Naïve classifier is briefly described as follows:

5.1.1 Naïve Bayes:

Naïve Bayes is an effective classifier when the outcome is known and the risk factors are conditionally independent. Key advantage is that the independence assumptions enables parameters for each variable to be learned separately, which is especially beneficial when dealing with large number of variables. It also gives the probability to determine a likelihood of the risk and to quantify the degree of uncertainty. (Wolfson et al., 2015) suggests that the Bayes classifier is reliable and maintains its effectiveness even when some data is missing or when dealing with datasets that have a large number of variables (high dimensionality).

- Naïve Bayes algorithm for binary outcomes: Let's consider the task to estimate the likelihood of an event *E* occurring within a fixed time period $(0, \tau)$. This estimation is based on individual characteristics $\boldsymbol{X} = (X1, \ldots, Xp)$, measured at a defined BL time $t = 0$. For $\boldsymbol{n}$ subjects who have their characteristics $\boldsymbol{X}$ measured at $t = 0$ are followed for τ time units to observe the occurrence of event E ($E = 1 \Rightarrow$ indicate event occurrence). The estimation is $P(E = 1 \mid X)$ (Wolfson et al., 2015).

$$P(E = 1 \mid X) = \frac{P(X \mid E = 1)P(E = 1)}{P(X)} = \frac{P(X \mid E = 1)P(E = 1)}{\sum_{e=0,1} P(X \mid E = e)P(E = e)} \tag{5.1}$$

- Naïve Bayes for Multiple outcomes: The equation is rewritten as

$$P(E = 1 \mid X) = \frac{\prod_{j=1}^{p} P(X_j \mid E = 1)P(E = 1)}{\sum_{e=0,1} \prod_{j=1}^{p} P(X_j \mid E = e)P(E = e)} \tag{5.2}$$

5.1.2 Confusion Matrix

A confusion matrix is used to evaluate the performance of classification algorithms. From the matrix the *True Positive, False Positive, True Negative and False Negative* can be determine. For example, see Figure 5.1. Section 1.2 provides a better understanding with numbers.

		Predicted	
		Good	**Bad**
Real	**Good**	True Positives (TP)	False Negatives (FN)
	Bad	False Positives (FP)	True Negatives (TN)

Figure 5.2: Structure of confusion matrix

With heart failure:

- *True Positive (TP)*: When a patient is correctly classified as an out-patient when they are indeed an out-patient, this result is termed as a true positive.
- *False Negative (FN)*: When a patient is incorrectly classified as dead when they are actually an out-patient, this outcome is termed as a false negative, also known as "***Type II error"***.
- *False Positive (FP)*: When a patient is incorrectly classified as an out-patient when they are actually dead, this classification is termed as a false positive, also known as "***Type I error/p-value***".
- *True Negative (TN)*: Finally, if all remaining patients are correctly classified as not alive (dead), then this classification is termed as true negative.

More advanced version of performance measures are:

- *Accuracy:* Accuracy assessment aims to quantify the effectiveness of the classifier's performance. An overall measure of classification accuracy can be derived from the confusion matrix by dividing by the total number of observations. In the context of heart failure, where timely and accurate diagnosis is critical for patient management and treatment decisions, accuracy ensures that the diagnostic results or predictions align closely with the actual clinical outcomes *(Writing Group et al.)*. However, there is a

drawback of this measure which is that it does not provide insight into how well individual classes are classified.

$$Accuracy = \frac{(TP + TN)}{(TP + FN + FP + TN)} \tag{5.3}$$

- *Sensitivity or true positive rate (TPR):* Sensitivity is determined by dividing the number of true positive instances by the total number of actual positive cases. It represents the percentage of actual positive observations that were correctly predicted (Alejandro & Banco, 2012). For example, in the context of heart failure diagnosis, high sensitivity ensures that individuals who truly have the condition are not missed during screening or diagnostic processes. This is crucial for timely intervention, treatment, and management of the disease, ultimately resulting in improved patient outcome and lower morbidity and mortality rates. Sensitivity statistic is computed as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{5.4}$$

- *Specificity or True negative rate (TRN):* Heart failure diagnosis involves distinguishing between patients who have the condition and those who do not. Specificity measures the proportion of individuals without heart failure correctly identified as such. In other words, it quantifies the ability of a diagnostic test or predictive model to classify healthy individuals as negative for heart failure correctly. High specificity indicates a low rate of false positives, which is crucial for ensuring that healthy individuals are not misdiagnosed or unnecessarily subjected to further tests or treatments:

$$Specificity = \frac{TN}{TN + FP} \tag{5.5}$$

where TN denotes the true negatives and FP denotes false positives.

- *Strength:* Strength refers to the total number of examples that are correctly classified by the rule during the training process. Strength is essential for heart failure clinical applications because it ensures the reliability of risk stratification, optimizes patient care, facilitates efficient resource allocation and provides valuable prognostic information.
- *Balanced Accuracy:* Balanced accuracy provides a balanced assessment of the classifier's performance across different classes, particularly in situations where there is class imbalance in the dataset (Wasikowski & Chen, 2010; Saito & Rehmsmeier, 2015; Buda et al., 2018). In heart failure datasets, the number of patients in different risk categories (e.g., alive, hospitalized, deceased) may not be evenly distributed, leading to class imbalance. Balanced accuracy takes into account the sensitivity and specificity of the classifier for each class, ensuring that the performance metrics are not biased towards the majority class. This is particularly important because misclassification of minority classes (e.g., deceased patients) can have significant clinical implications.

$$Balanced\ Accuracy\ (BA) = \frac{Sensitivity + Specificity}{2} \tag{5.6}$$

5.2 Evaluation

Naïve Bayes, Bayes Net and RBF Network classifier are applied to the HLL dataset to develop a model to classify patients into two different predefined risks.

5.2.1 Experimental setting

From the dataset described in chapter 4, I used following patient selection criteria for this analysis: All patients included in the study with an NT-proBNP measurement recorded at the baseline and remained alive for at least 8 months after the baseline visit OPD. This duration was chosen to ensure that sufficient time was allowed to observe the progression of their condition and to determine their survival status accurately. To be included in the analysis, patients must have had complete records for at least three of the following key clinical variables: Sodium (mmol/L), Potassium (mmol/L), Urea (mmol/L), Creatinine (mmol/L), Albumin (g/L), NT-proBNP. Weka tool is used for analysis.

This criteria resulted in a smaller dataset containing 1867 baseline records. From this dataset 204 records were randomly selected (using random number tables method) for the purpose of model training (test data). The test data has total 1663 records, consists 1645 alive and 14 deaths within between 8 to 12 months. Two different classifications schemes were used:

- Two classes - Dead and alive
- Three classes - Dead, hospitalization and out-patient visits.

5.2.2 Results

Let's start by examining a simpler classification scenario - 2×2 confusion matrix where risk is classified into two binary classes: dead and alive. This results in four sub-metrics namely: True Positive, False Positive, True Negative and False Negative, as shown in Table 5.1. Three different classifier (Naïve Bayes, Bayes Net and RBF Network) are employed for initial classification.

Baseline (n = 1663) – Alive/Dead						
Naïve Bayes	**Real**	**Predicted**		**Total**	**Performance**	**%**
		Alive	Dead		SEN	98.9
	Alive	1631	18	1649	SPEC	42.9
	Dead	8	6	14	ACC	98
					Balance Accuracy	70.9
					F1-Score	99.2
Bayes Net	**Real**	**Predicted**		**Total**	**Performance**	**%**
		Alive	Dead		SEN	99.6
	Alive	1643	6	1649	SPEC	71.4
	Dead	4	10	14	ACC	99.4
					Balance Accuracy	85.5
					F1-Score	99.7
RBF Network	**Real**	**Predicted**		**Total**	**Performance**	**%**
		Alive	Dead		SEN	99.9
	Alive	1648	1	1649	SPEC	14.3
	Dead	12	2	14	ACC	99.2
					Balance Accuracy	85.5
					F1-Score	99.6

Table 5.1: Binary Performance of the three classifiers using HLL dataset

From Table 5.1, it can be seen that Naive Bayes classifier demonstrates a high ability to correctly identify individuals who are truly alive with a sensitivity of 99% (i.e., correctly identifies 1631 out of 1649 alive patients and predicts only 18 alive patients as dead). The low error rates demonstrate strong performance in classifying positive observations. The specificity is lower at 43% which shows higher rate of false positives where individuals are classified as alive (8 out of 14) when they are actually dead.

The higher accuracy (98%) compared to the balanced accuracy (70.9%) indicates and imbalance in the dataset, Balanced accuracy is calculated as the arithmetic mean of sensitivity and specificity. It is particularly useful when dealing with imbalanced data, where one target class is significantly more frequent than the other.

The reasons for above are:

- Imbalanced dataset: The dataset might have unequal distributions among the classes. For example, in medical diagnostics, the number of healthy patients might vastly outnumber the number of patients with a particular disease. In this case, accuracy alone might be misleading because a classifier could attain high accuracy by consistently predicting the majority class for every instance.
- Impact of imbalance: When the dataset is imbalanced, accuracy can be inflated because the classifier might predominantly predict the majority class, which results in a high number of true negatives. However, this doesn't necessarily reflect the classifier's ability to correctly identify instances of the minority class.
- Feature imbalance: The features used for classification may not be equally informative for predicting both alive and dead classes. Certain features may be more strongly associated with one class than the other, leading to disparities in sensitivity and specificity.
- Importance of Balanced Accuracy: Balanced accuracy considers the sensitivity (recall) across all classes, giving equal weight to each class. This metric provides a more nuanced estimate of the classifier's performance, particularly in situations where class imbalance exists. A lower balanced accuracy suggests that the classifier struggles to identify instances from minority classes correctly.

You might be wondering about the distinction between Balanced Accuracy and the F1-Score, especially when dealing with imbalanced classification tasks. Both Balanced Accuracy and the F1-Score are valuable for evaluating models on imbalanced datasets, their suitability depends on the specific goals and priorities of the classification task. Balanced Accuracy excels in scenarios where equal attention is needed for both positive and negative outcomes, whereas the F1-Score is preferred when the emphasis is on correctly identifying positive instances while controlling false positives.

Bayes Net and RBF Network classifiers may exhibit similar patterns in sensitivity, specificity, accuracy, balanced accuracy and F1-score. However, the actual values may vary based on the characteristics of the classifiers and the dataset.

After analysing the BL to 12 months classification, I now look at 4 to 12 months classification. For this a subset of the dataset (used above) comprising 986 individuals who have had a 4-month's OPD follow-up visit. My objective is to classify these individuals at the 12-month mark. This approach allows us to assess the performance of the model using a new set of feature (different covariate values). This allows us to evaluate how the model adapts to and performs with this evolving dataset. For example, if certain health conditions become more prevalent in the dataset over time or if new treatment methods are introduced, these changes can shift the distribution and relationships of the features within the data. This enables us to understand the dynamic nature of patient health trajectories in a way. This insight is particularly valuable for chronic conditions where the disease progression can vary significantly among patients and over time.

Baseline (n = 986) – Alive/Dead					
	Real	**Predicted**	**Total**	**Performance**	**%**

Naïve Bayes		Alive	Dead		SEN	98.9
	Alive	961	11	972	SPEC	35.7
	Dead	9	5	14	ACC	98
					Balance Accuracy	67.3
					F1-Score	99

Table 5.2: Naïve Bayes performance of the patients at 1st event

In summary, Naïve Bayes (in Table 5.1) has higher specificity (43%) a slightly better ability to correctly classify individuals as 'Dead' when they are actually 'Dead', compared to (Table 5.2) Naïve Bayes (36%). Similarly, Naïve Bayes (Table 5.1) also demonstrates a higher balanced accuracy (70.9%) than (Table 5.2) Naïve Bayes (67.3%). This indicates that Naïve Bayes with bigger dataset and less missing features achieves better overall accuracy across both the 'Alive' and 'Dead' classes compared to Naïve Bayes smaller dataset. This also shows that there is imbalance in the binary classes.

To investigate the discrepancies further which are observed in Table 5.1 and Table 5.2, I have introduced another class to mitigate the impact of class imbalance on the binary classification. The aim was to determine whether making adjustments to the model could improve its performance by achieving a more balanced representation across the different categories of patient outcomes.

Table 5.3 illustrate the 3×3 confusion matrix. I used the same dataset as for binary classification, but 'Alive' patients were further categorized into more informative groups: 'Hospitalization' (patients who had been admitted to the hospital in the last 12 months) and 'Out-patient' (patients who had visited the heart failure (HLL) clinic but were not admitted to the hospital in the last 24 months).

Table 5.3 presents the model performance. Out of 1663 individuals, there were 663 hospitalizations, 14 deaths, and 986 OPD instances. The classification model was retrained using the same 204 patients, but with consideration of three classes. In the matrix below, the predicted classes are displayed in the columns, while the actual classes are in the rows.

Baseline (n = 1663)					
Classifier	**Real**	**Prediction**			**Total**
		Hospitalisation	Dead	Out-patient	
Naïve Bayes	Hospitalisation	422	5	236	663
	Dead	3	6	5	14
	Out-patient	85	8	893	986
Bayes Net	Hospitalisation	478	2	183	663
	Dead	1	11	2	14
	Out-patient	102	4	880	986
RBF Network	Hospitalisation	502	1	160	663
	Dead	6	3	5	14
	Out-patient	89	0	897	986

Table 5.3: Multi-class risk classification for the patients with HF

The first matrix is for Naïve Bayes classifier, shows 422 events of hospitalisation, 5 event of death and 893 out-patient individuals were correctly classified (diagonal row). For actual hospitalisation (n = 663), 422 instances were correctly classified as hospitalisation, while 5 were misclassified as Dead and 236 as Out-patient. Similarly, Dead and Out-patient can be interpreted.

The table also provides the performance of other classifiers Bayes Net and RBF Net. This shows accuracy and reliability in classification of patient s across three classifier.

As discussed earlier, observing the performance of multi-class classification model - four matrices namely; TR, FP, TN and FN for each class are need to be assessed. Example is given for Naïve Bayes classifier. Using table 5.3 data I calculate the following:

- *True Positive (TP):* The number of instances correctly classified as positive. I.e., (Hospitalisation (n = 422), Dead (n = 6) and Out-patient (n = 893)).
- *False Negative (FN):* The number of instances incorrectly classified as negative (sum of values in the corresponding row, excluding the TP). I.e., Hospitalisation (5 + 236 = 241), Dead (3 + 5 =8) and OPD (85 + 8 = 93).
- *False positive (FP):* The number of instances incorrectly classified as positive when they are actually negative (sum of values in the corresponding column, excluding the TP). I.e., Hospitalisation (3 + 85 = 88), Dead (5 + 8 = 13) and OPD (5 + 236 = 241),
- *True Negative (TN):* The number of instances correctly classified as negative (sum of all columns and rows, excluding that class's column and row). I.e., Hospitalisation (6 + 5 + 8 + 893 = 912), Dead (422 + 85 + 236 + 893 = 1636) and (422 + 5 + 3 + 6 = 436).

Baseline (n=1658)											
Classifier	**Classes/Outcome**	**TP**	**TN**	**FP**	**FN**	**Sensitivity (TRP)**	**Specificity (SPEC)**	**Accuracy per class**	**Accuracy classifier**	**Balanced Accuracy**	**F1 Score**
Naïve Bayes	Hospitalisation	422	912	88	241	63.7	91.2	80.2	79.43	66	72.0
	Dead	6	1636	13	8	42.9	99.2	98.7			36.4
	Out-patient	893	436	241	93	90.6	64.4	79.9			84.2
Bayes Net	Hospitalisation	478	897	103	185	72.1	89.7	82.6	82.3	80	76.8
	Dead	11	1643	6	3	78.6	99.6	99.5			71.0
	Out-patient	880	492	185	106	89.3	72.7	82.4			85.8
RBF Network	Hospitalisation	502	905	95	161	75.7	90.5	84.6	84.31	63	79.7
	Dead	3	1648	1	11	21.4	99.9	99.3			33.3
	Out-patient	897	512	165	89	91.0	75.6	84.7			87.6

Table 5.4: Performance of multi-class risk classification for the patients with HF

Based on the calculated metrics, the performance comparison of the three classifiers:

- Naïve Bayes:

Sensitivity (Recall) and specificity values for Naïve Bayes are moderate across all classes. Accuracy varies across classes, with the highest accuracy observed for the Dead class (98.7%) and the lowest for the Out-patient class (79.9%).

- Bayes Net:

Bayes Net exhibits higher sensitivity and specificity values compared to Naïve Bayes across all classes. Accuracy per class is generally higher for Bayes Net compared to Naïve Bayes, with the highest accuracy observed for the Dead class (99.5%).

- RBF Network:

RBF Network shows comparable sensitivity and specificity values to Bayes Net, with slightly lower performance in some cases. Accuracy per class for RBF Network is generally similar to Bayes Net, with high accuracy observed for the Dead class (99.3%).

Comparing performance of binary (Table 5.1) and multiclass (Table 5.4) classification, I looked the results of Naïve Bayes classifier in both model. It can be seen that binary classifier has a high accuracy of about 98% while the multiclass classifier shows a lower overall accuracy 79%. The prediction of *Dead* in multiclass model is considerably low. This suggest challenges in identifying this class possibly due to fewer samples or overlapping features with other

classes. Table 5.4 shows for Naïve Bayes achieving higher accuracy - (79%) than balanced accuracy (66%) indicating residual imbalance in the dataset despite adding another class.

Similarly, as mentioned earlier in binary classification section where I have looked on 4 months to 12 months classification using the same model but here with 3 classes. Table 5.6 compared to Table 5.2 shows the accuracy has decreased in the multiclass model. Balanced accuracy shows that imbalance is still exists in the data and missing data affect the results.

OPD (986)					
Classifier	**Real**	**Prediction**			**Total**
		Hospitalisation	Dead	Out-patient	
Naïve Bayes	Hospitalisation	131	1	113	245
	Dead	2	6	6	14
	Out-patient	33	2	692	727

Table 5.5: Classification (OPD) within the 12 months of BL event

Those who were classified as OPD (986)											
Classifier	**Classes / Outcome**	**TP**	**TN**	**FP**	**FN**	**Sensitivity (TRP)**	**Specificity (SPEC)**	**Accuracy per class**	**Accuracy classifier**	**Balanced Accuracy**	**F1 Score**
Naïve Bayes	Hospitalisation	131	706	35	114	53.5	95.3	84.9	84.1	64	63.7
	Dead	6	969	3	8	42.9	99.7	98.9			52.2
	Out-patient	692	140	119	35	95.2	54.1	84.4			90.0

Table 5.6: Naïve Bayes performance – for table 5.5

5.2.3 Performance measure

As discussed in 5.2.2 section, identifying fatal cases consistently poses a challenge for both models. The risk classes within the dataset exhibit imbalance, with notably smaller number of deaths in the first 12 months compared to the number of patients who remain alive (those in hospital or OPD state). This imbalance poses challenges to a classifier's capability to learn to identify the dead class effectively. As the result, the confusion matrices that report the average precision and recall across risk classes likely to overestimate the true performance. Classifiers such as Naïve Bayes typically operate under the assumption that the training data is sufficiently large and that the continuous values associated with each class follow a Gaussian distribution. These assumptions are not met by the HLL dataset.

5.3 Conclusion

Data is from the single centre Researchers must interpret their results with caution, considering the implications of class imbalance and model assumptions. Notwithstanding, this chapter provides basic insights into supervised machine learning risk classification methods. As shown in the results that one of the main challenges facing the development of machine learning-based classifiers for clinical datasets is the large dimensionality and skewness of the data (which is shown in the chapter 4). This underscores the importance of considering balanced accuracy and other metrics to provide a detailed evaluation of classifier performance, particularly in the presence of class imbalance. Even though the introduction of another class into the dataset did not help to alleviate the imbalance. It's essential to recognise the potential

limitations stemming from dataset characteristics. For instance, in the context of heart failure studies where the number of deceased patients may be significantly lower compared to those who are alive or hospitalized. Introducing a balanced class for deceased patients can help mitigate this imbalance, but it's crucial to do so carefully to ensure the new class accurately reflects the data distribution without introducing bias or noise. Researchers should also validate the classifier's performance on the augmented dataset to ensure it effectively addresses the imbalance issue without compromising the model's ability to generalize. Exploring dynamic risk models may further improve the model's performance and overcome challenges posed by dataset characteristics.

Classification is useful for organizing and categorizing data into discrete classes based on predefined features. It has limitations when it comes to dynamic risk analysis, which involves predicting future events or outcomes based on changing variables and conditions. There are number of reasons why these methods may not be giving as good predictions. Some of these are:

Assumption of independence: Naïve Bayes, for example, assumes that features are independent of each other given the class label. This simplifying assumption may not hold true in real-world scenarios (Rigby, 1991) where variables are interconnected and influence each other's behaviour over time. Dynamic risk analysis can capture these dependencies and provide more accurate predictions.

Inability to adapt: Classification models are trained on historical data and may not adapt well to changing conditions or new information. They do not have the capability to update their predictions in real-time as new data becomes available. Dynamic risk analysis, on the other hand, can incorporate new data and adjust risk assessments accordingly.

Risk Assessment vs. Prediction: While classification methods are effective for predicting class labels based on input features, they may not provide insights into the underlying risk factors or mechanisms driving those predictions. Dynamic risk analysis goes beyond prediction to assess the probability and impact of different risk factors over time, helping stakeholders make more informed decisions.

It's important to note that classification alone doesn't provide insights into the individual prognosis of each patient. To truly understand the prognosis for HF patients, I may need more advanced predictive methods especially machine learning techniques that typically falls under the domain of clustering or unsupervised learning. Given these drawbacks and disadvantages, as will be seen in the following chapters using time dependency dynamic risk modelling is more suitable for predicting the progression of patients.

Chapter 6 Dynamic risk modelling using Absorbing Markov chains: Short and long term prediction

In this chapter, I look at dynamic risk modelling and how to apply it to the HLL. The HLL is detailed in chapter 4. The aim is to understand the progression patterns and probabilities of different health outcomes in patients with HF. This help them make informed decisions and manage their care effectively. Figure 6.1 illustrates the various domains that are related to a patient's prognosis, focusing on both the quantity and quality of life.

Figure 6.1: Domains that are related to a patient's prognosis

- focusing on both the quantity and quality of life. (Allen et al., 2012)

As discussed in the introduction chapter, for patients with CHF, the clinical interest focuses not only in the final outcome but also on the dynamics of disease progression, mainly concerning the requirement for hospitalisation (Khand et al., 2001; Ieva et al., 2017; Jiang et al., 2019). A robust model might support prediction at the individual patient level while also providing estimates of the health care resources required to meet patient needs.

In managing health care needs, current models focus on the flow of patients between available resources (such as staff and departments along a pathway of care). Models do not focus on **when** and **how often** the care provision is required using the current and previous condition of the patients. Changes in the patient's condition will lead to a change in the clinician's judgement of prognosis. The care of patients requires frequent reappraisal of their clinical trajectory, which helps to calibrate risk in a dynamic fashion. In the case of HF, the assessment of prognosis is not straightforward and is in contrast to the more linear decline of a patient with advanced cancer (Uhry et al., 2010). The dramatic variation in disease severity and its unpredictability makes decision-making extremely difficult.

Markov Models, incorporating the rate of change in multiple indices[2], can potentially forecast adverse events (hospitalisation and death) (Ieva et al., 2017). These are particularly useful in studying of chronic illness, as they categorise a patient's condition into a finite number of distinct states at any given point in their clinical trajectory. In this chapter, I use Markov models to predict disease progression in heart failure and thus allow for a better understanding of the flow of patients between different states of health.

To model the progression of HF, I therefore used Markov chains in a well-characterised[3] cohort of patients referred for investigation of potential HF. This approach is based on a finite number of mutually *exclusive* and *exhaustive* distinct states. The following analysis is carried out to see if I could use events early in a patient's career to predict their

[2] Multiple indices refers to a variety of measures or indicators that are used to assess different aspects of a condition or situation.

[3] A well-characterized cohort means patients are fully documented with clinical signs, symptoms, and all relevant data for heart failure diagnosis.

likely outcomes to them during subsequent FU. This could enhance insight of the patient pathways through HF and facilitating more effective service planning and resource distribution.

Further details and an in-depth discussion of the Markov model are provided in the subsequent section of this chapter. The UK's National Institute for Health and Care Excellence (NICE) uses various methods and models for health economic evaluations. One of the commonly used methods is the Markov model (NICE, 2018a).

6.1 Heart failure risk assessment

Over time, the changes in the patient's condition can cause a change in the prediction of risk, which results in a revision of the care and treatment strategy. As suggested earlier, I am interested in determining the timing (when) and frequency (how often) of required patient care. This results in a number of features that are to do with the cost associated, survival and quality of life.

Figure 6.2 presents the components influenced by risk assessment in clinical practice, connecting it to actions like effective triage[4] and specialist referrals, while also steering the allocation of resources and informing patients. This process is integral to tailoring healthcare interventions, as depicted in Figure 6.1, which resulted in a number of features that have to do with the cost associated, survival and quality of life.

[4] Effective triage is a critical process in healthcare settings used to prioritize patients based on the urgency of their need for care. It involves assessing the severity of patients' conditions to determine who requires immediate attention and who can wait for care.

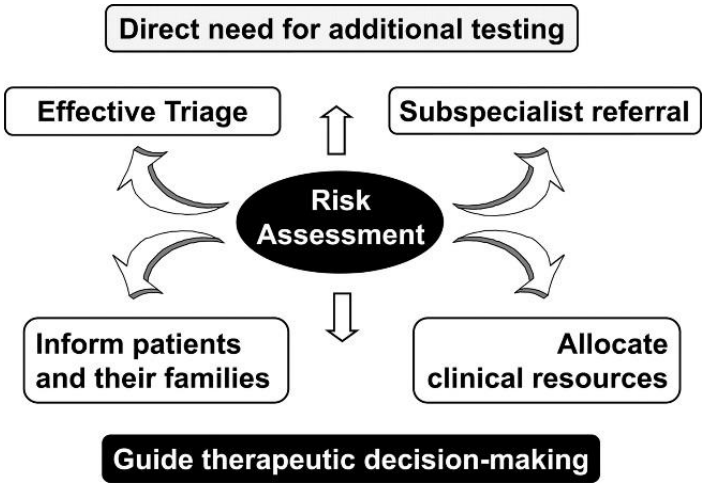


Figure 6.2: Goals of risk assessment in patients with CV disease

6.2 Markov Models and Chains

In chapter 3, section 3.5.1, some elements of Markov chains have already been introduced. In this section, I will discuss about it in more detail. A *stochastic process* refers to any process that that evolve with uncertainty and models based on these processes are referred to as stochastic or probabilistic models (Sonnenberg & Beck, 1993; Mhoon et al., 2010; Sato & Zouain, 2010). Especially, the event-based progression (and other complexities of a disease) can be represented using multistate models (Andersen & Keiding, 2002). These are often based on the use of first-order Markov processes (chain and hidden) and allow the risk to evolve dynamically.

In multi-state models for intermittently observed processes[5], the times at which state changes occur are usually known to be within bounded or fixed time intervals, making them interval-censored. Markov models represent stochastic processes that evolve over time and consist of a finite number of states, states which change based on transition probabilities. If the process is stochastic and the state or behavior of model at any given time period (i.e., cycle) does not depend on its state or behavior in previous time period then the process is *Markovian*. The future state depends only by the present state independent of events that occurred in previous states. Hence:

- The process has "lack of memory" (Singer et al., 2014; Wu & Chu, 2017; Schröger et al., 2023). That is current state of the model determines what state the model can change to next time point. The trajectory on how current state emerged does not matter; in this sense it is memoryless.
- Even processes where the previous state is relevant can be converted to Markovian by defining temporary states, known as *tunnel states.*

A Markov process in which the transition between states is based on constant probabilities is called a Markov chain (Sonnenberg & Beck, 1993). The major advantage of Markov Chain Modelling (MCM) is its flexibility to model complicated events and related situations for which algebraic solutions are not possible. Such modelling is more applicable when interest lies in estimating either the likelihood of transitioning from one state to another within a specified time period, or the average time spent in a state (mean sojourn time). When applied to illness,

[5] Intermittently observed processes" refer to processes or events that are not continuously monitored or recorded but are instead observed at discrete and irregular intervals. In this analysis, we used the underlying disease progression as continuous, with clinical events captured at discrete time points. This approach enables us to approximate the continuous nature of disease dynamics within the framework of discrete observations, facilitating a comprehensive analysis of patient outcomes.

the Markov model assumes that a patient can be in any one of the many states referred to as '*Markov states*' and that there are '*events*' which allow the transition from one state to another within a specified time period, the process known as a "*Markov cycle"* (Sonnenberg & Beck, 1993). The memory-less property enable model to be represented using a single-cycle transition matrix. A regular Markov chains (RMC) are governed by definition, where:

A *stochastic process* is a MC if for all times $n \geq 0$ and all states $i_0, \cdots, i, j \in S$,

$$\mathcal{P}\left(X_{n+1} = j \middle| X_n = i, X_{n-1} = i_{n-1,}, \ldots, X_0 = i_0\right) = \mathcal{P}(X_{n+1} = j | X_n = i,) \quad (6.1)$$
$$= p_{ij} \quad \forall\, i \leq n, j \leq n$$

X represent the stochastic process, specifically the sequence of random variable that denote the state of the process at each time step. For example, X_n is the state of the process at time n. S denote the state space of the MC. It is set of all possible states that the process can take. For instance, if $S = \{0, 1, 2\}$, the process can only be in one of these three states at any given time.

$\mathcal{P}_{ij}$ is donating the probability that the chain when in state i, moves to the next state j one unit of time, and is often called to as a "*one-step transition probability*". The square matrix

$$P = (\mathcal{P}_{ij}), i, j \in S \quad (6.2)$$

is referred as a one-step transition matrix. Since the chain must transition from state i, to one of the state j, each row must sum to 1, i.e.,

$$\sum_{j \in S} \mathcal{P}_{ij} = 1 \quad (6.3)$$

If I make an assumption that the transition probabilities do not depend on "n" (the time), then using $n=0$ in (6.1) gives

$$\mathcal{P}_{ij} = P(X_1 = j | X_0 = i) \tag{6.4}$$

The key property of the chain is, that the next future state is dependent given the present state irrespective of past state. Therefore if n is the present time, then the future is given by $\{X_{n+1}, X_{n+2}, X_{n+3} \ldots . X_{n+m}\}$ while the past of the chain is given by $\{X_0, X_1 \ldots . X_{n-1}\}$ and current state is $\{X_n\}$. The matrix containing $\mathcal{P}_{ij}$, the transition probabilities of n states, can be represented as a $n \times n$ matrix (P) as shown below.

$$\begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n,1} & p_{n,2} & \cdots & p_{n,n} \end{bmatrix} = \mathcal{P} \tag{6.5}$$

where $p_{i,j}$ represent the probability (p) of transition from i (starting state) to j (next state of immediate transition). Eg.,, in equation (a) $p_{1,2}$ denotes the probability of transitioning from state s_1 to state s_2. Likewise, the probability of transitioning from s_2 to s_1 is given by $p_{2,1}$. Note that $p_{1,2}$ is not necessarily the same as $p_{2,1}$.

Regular Markov Chain models have two important properties: *irreducibility* and *aperiodicity*. Looking at the 2 states' (i.e., i and j) model, if state j is $accessible$ from state i, $i \rightarrow j$ then if $P_{ij}^n > 0$ for some $n \geq 0$, implies that j can be reached from i in finite number of steps. If j is not $accessible$ from i then $P_{ij}^n = 0$ for all $n \geq 0$, and thus the chain started from i never visits j. Whereas, if i is accessible from j, and j is accessible from i then this

means i and j communicate with each other and can be represented by $i \leftrightarrow j$. The accessibility relation groups states into distinct $class$. Within each class, every state can transition to every other state within the same class (they "communicate" with each other). However, states in different classes cannot transition to each other (they do not "communicate"). Essentially, there is full connectivity within each class, but no connectivity between states in different classes. A regular Markov chain is considered *irreducible* if all states are part of a one class, meaning that every state can communicate with each other. For example, if the chain has *n* states, *irreducibility* implies that the entries of matrix sum $I + P + P^2 \ldots + P^n$ are nonzero. This matrix sum represents the cumulative probabilities of transitioning from any state to any other state within *n* steps, including staying in the same state (where I is the identity matrix and P is the transition matrix. Since the element P_{ij}^n are between 0 and 1 (i.e., $0 < P_{ij}^n < 1$), for irreducible states the equation (1) can also be expressed as:

$$\forall i, j \in S, \exists m < \infty : P(Xn + m = j \mid Xn = i) > 0 \qquad (6.6)$$

Also, $irreducibility$ property of the chain force either all of the states are $transient$ or all are $recurrent$. A state i is said to be $transient$ if,

- Once the process enters state i, there is a positive probability that it may never return to state i again.
- there is another state j (distinct from i) that can be reached from i but i cannot be reached from j.
- In a finite-state Markov chain, transient state can be visited only a finite number of times.

A state i is considered $recurrent$ if,

- Once the process enters state i, it is certain to return to state i eventually.
- Because a $recurrent$ state is guaranteed to be revisited after each visit, it needs to be visited an infinite number of times.

In the case of aperiodicity, there is no return to the state i after entering to state j, so at this point, the state j has started keeping communication to itself only. Secondly, to check the behaviour of the Markov chain until absorption. For example the average time spent in any other state $j \neq i$ before being stuck in state i. Such a limitation of a regular Markov chain can be handled by a special type of state, one which when the process enters it, it never leaves it. Such states are referred to as $absorbing\ states$.

6.3 **Absorbing Markov chain**

In absorbing chain modelling (ACM), the terminal or censored states are fixed points or steady states: once a patient enters one of these states, there is no exit from it. Death is the primary example of such a state. However, there are also not-terminal states where individuals may reside without reaching an absorbing states for the duration of the study. This happens when a subject is alive and the study is ongoing, or the individual has withdrawn or been lost to follow-up. In the context of AMC is, an absorbing state is defined as follows:

When a chain enters a state X_A and remains in that state indefinitely, the probability of transitioning from this state to any other is zero, and the probability of remaining in the same state is one:

$$P(X_A) = \mathcal{P}_{00} = 1 \tag{6.7}$$

Thus X_A is called an absorbing state.

If a Markov chain reaches a steady state, the probabilities in every cycle remain same. This means that regardless of the number transitions *n*, the transition matrix raised to the power of *n* will produce the same probabilities. Therefore, if P_1 is the transition matrix after one cycle, then for any number of cycles *n*, the transition matrix P_n can be represented as:

$$P_n = P_1^n \tag{6.8}$$

However it is crucial to understand that:

- Steady-state predictions are never achieved in reality due to a
  - errors in estimating the transition matrix *P*
  - variation in *P* over time
  - shifts in how the states depend on each other - nature of dependency in relationships among the states.
- Steady-state probabilities may only exist if Markov chain is ergodic[6].

[6] Steady-state probabilities represent the long-term behaviour of the Markov chain, may not be defined or achievable unless the Markov chain is ergodic. An ergodic Markov chain is one where it is possible to reach any state from any other state in a finite number of steps, and all states are recurrent and aperiodic. Essentially, ergodicity ensures that the system will eventually explore all states sufficiently, allowing steady-state probabilities to exist.

The system can be modelled as an AMC by modifying the transition matrix (P) of RMC into a canonical form with four block matrices I (identity matrix for absorbing states), 0 (zero matrix indicating no return to transient states from absorbing states), R (transitions from transient states to absorbing states) and Q (transitions among transient states only) This structure is further detailed in section 6.5.2 (equation 6.9) where study data is used to demonstrate these concepts. It explains how these submatrices help to determine the *fundamental matrix* (equation 6.14) and the *limiting matrix* (equation 6.19).

6.4 **Model structure and specification**

The figure (6.3) presents a graphic representation of the study's methodology. It illustrate the number of patients with suspected heart failure referred to HLL clinic from 2000 to 2017 plus exclusion criteria. The detail of the HLL patients was provided in the chapter 4 section 4.4. Patients are categorised based on ESC guidelines, then transformation of patient data into distinct states for analysis, and the subsequent steps of AMC to predict the patient outcomes over short and long terms.

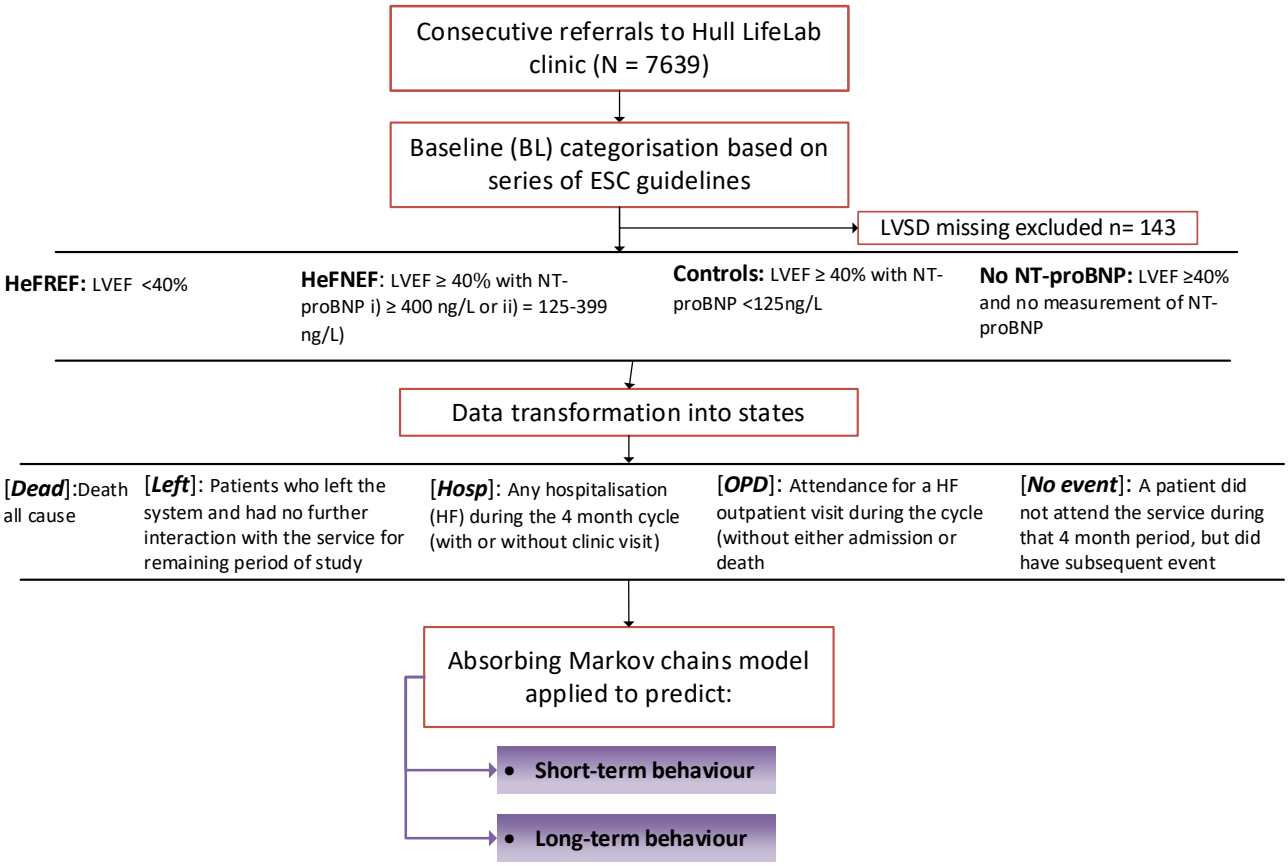


Figure 6.3: Graphical representation of the study's methodology

6.4.1 Selection of patients (their diagnostic categories and definitions)

Patients were categorised into different cohorts at baseline using definitions of heart failure based on a series of ESC guidelines (published between 2008 and 2016). Left ventricular systolic function (LVEF) assessed by echocardiography with different cut-offs for NT-proBNP (as defined by the NICE (NICE, 2010; 2018b) and European Society for Cardiology (McMurray et al., 2012; Ponikowski et al., 2016) were used to classify the population into different cohorts as follows:

1) HeFREF – those with LVEF < 40%.
2) HeFNEF – those with LVEF ≥ 40% and NT-proBNP:
    i) ≥400 ng/L (NICE, 2018b)
    ii) 125-399 ng/L (Ponikowski et al., 2016)
3) Controls – individuals who did not meet the criteria for cardiac dysfunction (i.e., criteria is defined as LVEF ≥ 40% and NT-proBNP < 125 ng/L).
4) No NT-proBNP – This group of patients had an uncertain diagnosis of HF (characterised by an LVEF ≥ 40% and no record of NT-proBNP).

NT-proBNP was introduced during the course of the study, this is not available for all patients. “Controls” are referred to in inverted commas patients because despite having normal cardiac investigations were not considered *normal* as a referring clinician suspected HF was a possible diagnosis. Patients without baseline LVEF were excluded from this analysis because their heart failure phenotype could not be classified (n=143; Figure 6.4).

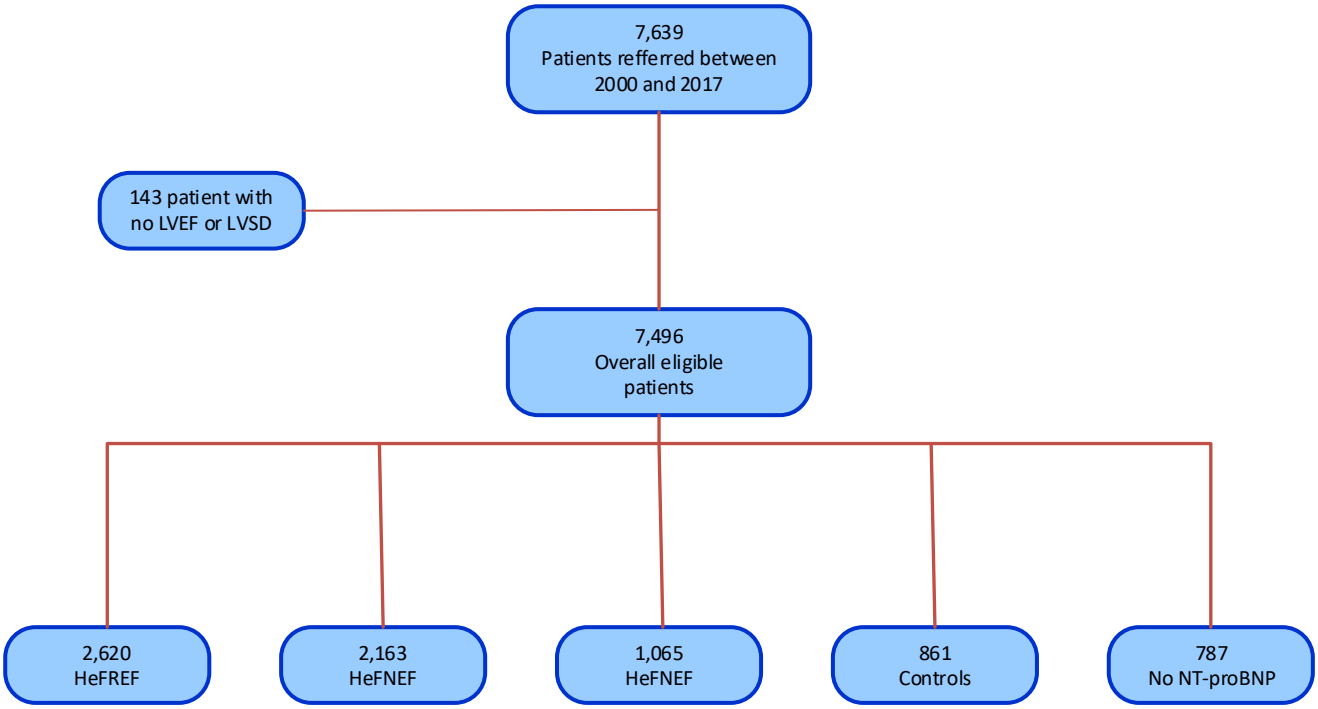


Figure 6.4: Consort diagram illustrated the flow of patients Patients with suspected heart failure referred to the HLL clinic from 2000 to 2017 their classification into different cohorts at BL.143 patients with no LVEF or LVSD recorded at baseline were excluded from the study

6.4.2 Data transformation and state definition

First, the data must be transformed into appropriate heath states after the diagnostic categorisation of the population. This allows to represent changes in the patient's health status over different risk phases (e.g. acute, unstable and stable) by a Markovian transition process. The patient's states were assessed at consecutive four-month intervals (cycles) from BL up to 24 months;

(1) $[Dead]$ – death from any cause.

(2) [$Left$] – patients who exited the system and had no further interaction with the service, neither died nor used the service for the rest of the study period.

(3) [$Hosp$] – any HF hospitalisation during the 4 month cycle regardless of a clinic visit;

(4) [OPD] – attendance for a HF out-patient visit during the interval - without admission or death.

(5) [$No\ event$] – a patient did not attend the service during that 4 month period, but did have a *subsequent* event, thus not classified as [$Left$].

[$Left$] and [$No\ event$] were considered as non-clinical Markovian states, indicating periods when the HF service was not utilised. This also help to complete the patient's history with no gap between any consecutive cycles. Patients were not excluded simply because no transition was seen during a particular time frame: it is possible for a patient to return from [$No\ event$] state to any of the states other than[$Left$].

The model allows the transition to [*Dead*] to occur at any time within a 4 monthly cycle. My approach was hierarchical: if a patient was both admitted and died within a single cycle, only the death is considered in the model. For instance, if a hospitalisation or death occurs within a cycle, the patient is assigned to that health state throughout that 4 monthly cycle. If patient have more than one OPD visit within the 4 monthly interval, the latest OPD visit is selected. Similarly the latest admission were considered. The underlying disease process is continuous, and clinical events are represented at discrete time points (Gruger et al., 1991; Jackson, 2019).

Figure 6.5 illustrates some possible transitions that patients with HF might follow in the 5 state model. It is extended form of Figure 4.1, where there were only 3 states.

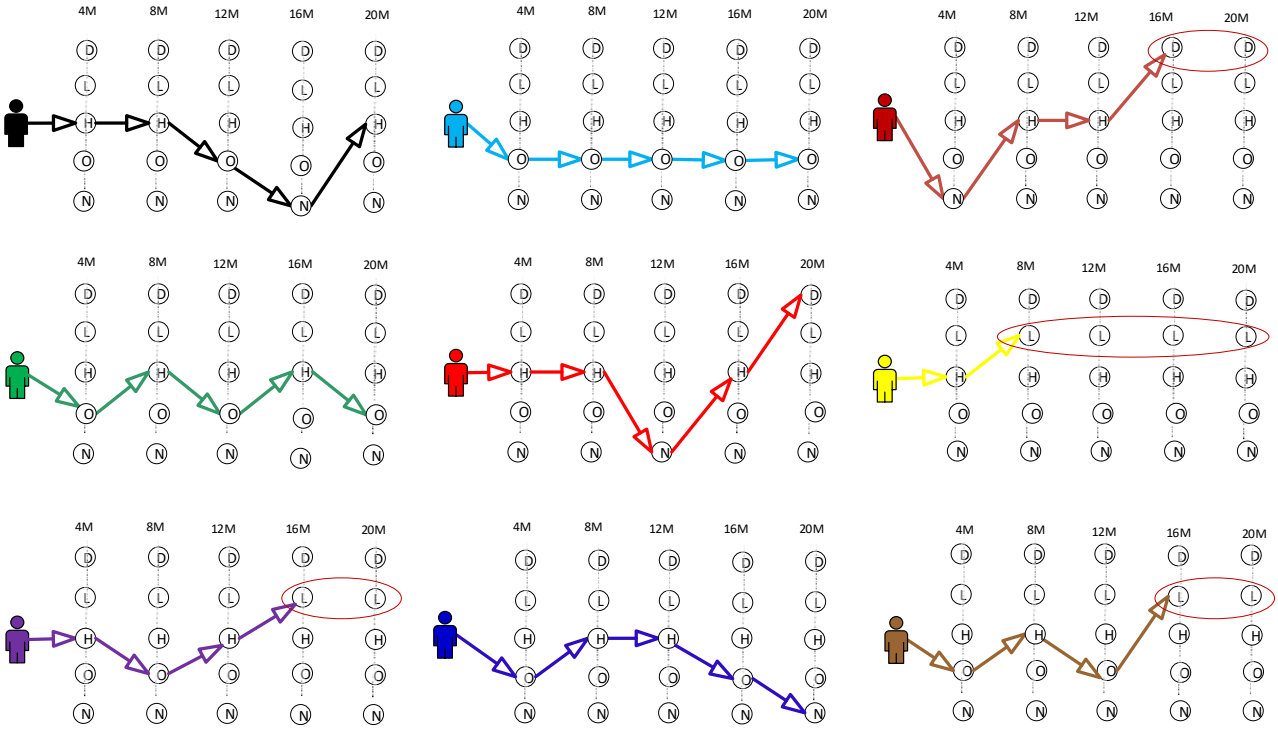


Figure 6.5: 9 cases of patient's healthcare journey for 5 states model.Figure illustrates potential transitions for patients up to 5 cycles after baseline. Abbreviations: D: [$Dead$],; L: [$Left$],; H: [$Hosp$],; O: [OPD], & N: [$No\ event$],. 4M: 4 month; 8M: 8 month; 12M: 12month; 16M: 16 month; and 20M: 20 month. An oval shape is used to indicate that individual transited to either [*Dead*] *or* [*Left*] state will stay in these states for the remaining cycles or until a process finished.

6.4.3 State transitions in the model

Given the 5 states, {[$Dead$], [$Left$], [$Hosp$], [OPD], [$No\ event$]}, the Markov model forecast the probability of individual being in specific states over time. E.g., if a patient is in the hospitalised state, what is the likelihood of either repeated hospitalisation or transitioning to another state subsequent intervals?

Transition matrices were created for each of the initial two transitions: (a) between BL to the end of 4 months (1st cycle) and (b) from 4 months (1st cycle to end of the 8months (2nd cycle) respectively). Since all patients starts in the $[OPD]$ state, the 1st transition is represented as a single line only (see Table 6.2). The study used these two matrices to capture and understand the dual temporal prospects, the immediate and extended risk patterns in patients with HF. The focus is on both short term (one-step probabilities through the sixth cycle (i.e., 24 months)) and long term (up to a maximum of 4 years) prediction. Figure 6.6 shows the follow of patients through different states tracked over two cycles. There are percentages showing the overall distributions of at the end of the follow-up cycle.

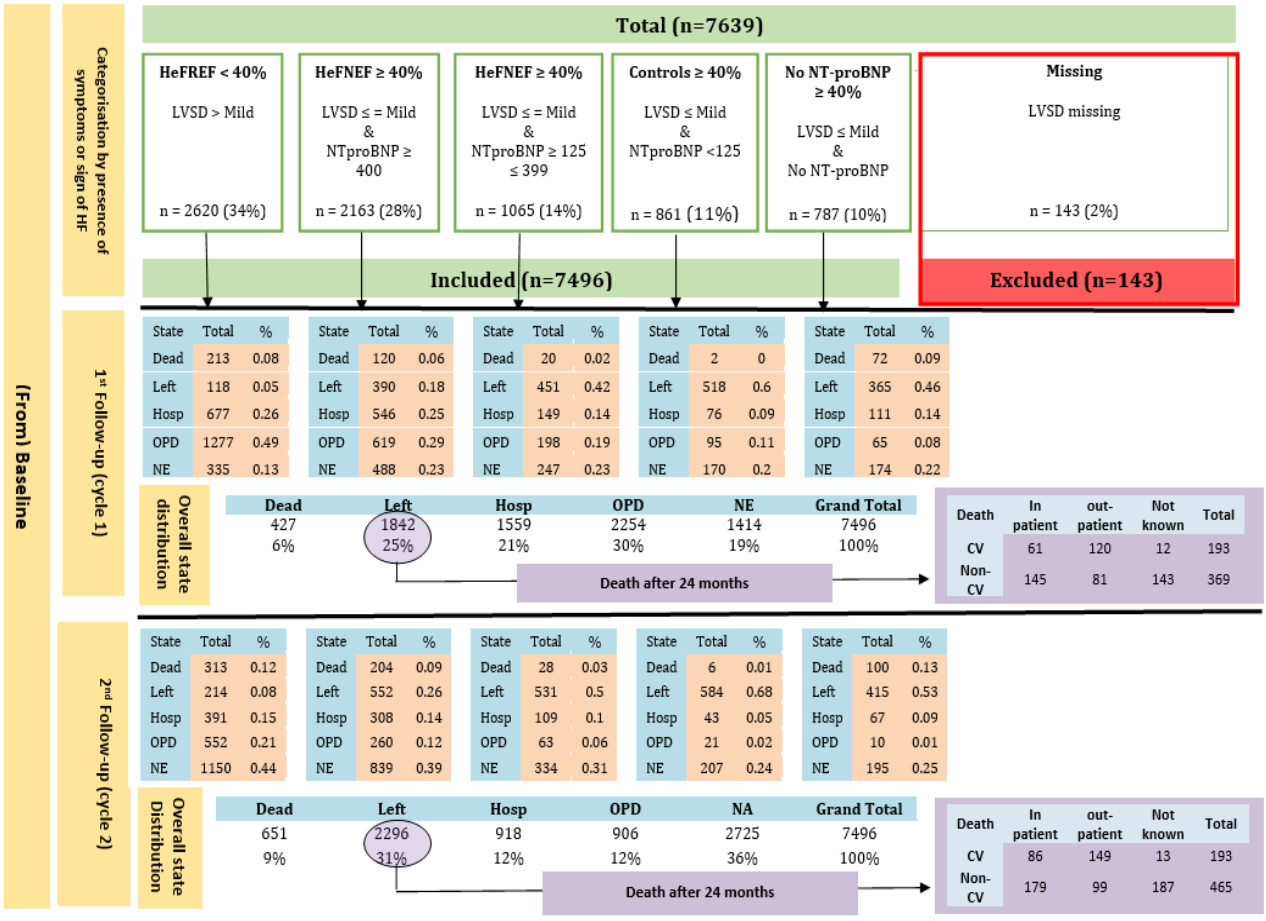


Figure 6.6: First two transitions among states as per diagnostic categories

Figure 6.7 illustrates the underlying five-state model used to examine disease progression. The directions of instantaneous transitions are indicated by the arrow. The transitions between transient states ($[Hosp]$, $[OPD]$ and$[No\ Event]$) are both direction but no further transition take place once the absorbing state is reached, no further transitions are possible. In this analysis there are two absorbing states ($[Dead]$ and $[Left]$). This figure is compact form of figure 6.5.

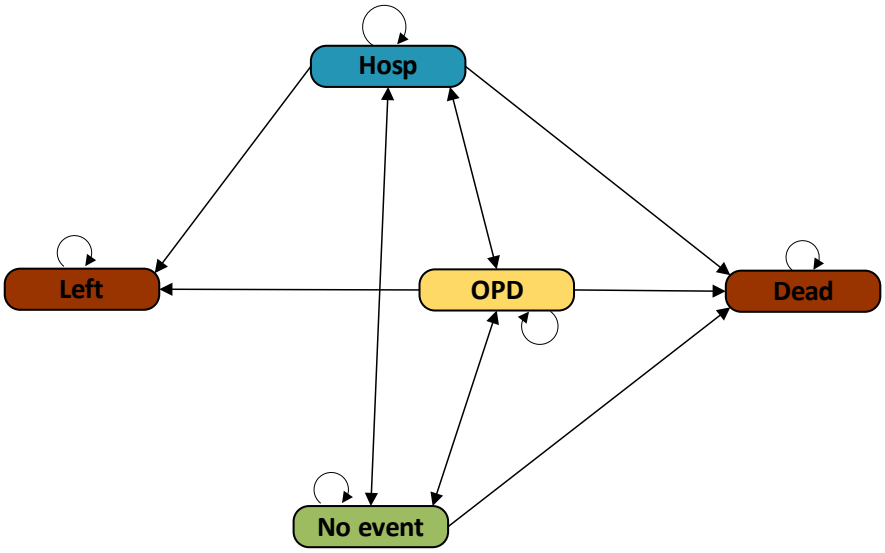


Figure 6.7: Underlying 5 state model for examining the HF disease progression

6.5 Results

The model was developed using the observed frequencies of state transitions from the first two cycles. I subsequently applied this model to the original dataset to predict future outcomes. To evaluate the model's accuracy, I compared the predicted transition probabilities with the observed transitions up to sixth cycle. For examining the model's long-term behaviour, I computed a *fundamental matrix* (6.14) and a *limiting matrix* (6.19) based on modelled data.

By incorporating dual temporal prospects, I analysed both the immediate and extended risk patterns in patients with HF. For examining the model's long-term behaviour, I computed two key matrices: the *fundamental matrix* (equation 6.14) and the *limiting matrix* (equation 6.19). The *fundamental* matrix provides information on how long an individual is likely to remain in each transient state and estimates how many cycles it will take before the individual reaches a permanent (absorbing) state within the model's duration. The *limiting* matrix shows the expected probabilities or proportions of individuals eventually reaching one of the permanent (absorbing) states. More detailed information is available in section 6.5.1.

6.5.1 Baseline demographics

6.5.1.1 Diagnostic categorisation

The study includes 7,496 patients. Patients are categorised as: 34%(N=2620) patients had HeFREF, 28%(N=2163) patients had HeFNEF with NT-proBNP ≥400 (ng/L), 14%(N=1065) pateints had HeFNEF with NT-proBNP between 125 & 399 (ng/L), 11%(861) referred as ("controls") because they did not meet the criteria for HF, 10%(N=787) patients had uncertainty when diagnostic categorisation, due to normal LV systolic function but no NT-proBNP result, referred to as "No NT-proBNP") (Figures 6.3). Table 6.1 provides a detailed overview of the demographic and clinical characteristics of patients in each diagnostic category.

Continuous variables are reported as median and interquartile range (IQR), while categorical variables are displayed as count and percentages. The independent *t-test* was used to evaluate differences in continuous data between the diagnostic groups. After transforming the data into health states, the distribution of patients across each state is presented in tables. *P*

values, calculated from analysis of variance, represent the differences between patients in diagnostic categories.

Various packages were used for analyses, including R (version 2022.02.1), Stata software, and Excel. Statistical significance was determined by a two-sided p-value of less than 0.05.

Variable	Missing (n)	Total	HeFREF	HePNEF		Control	No NT-proBNP	*P* value
				>=400	125-399			
Demographics								
Age (years)		73.5 (65.6, 79.9)	72.6 (64.4, 78.9)	77.3 (71.3, 82.8)	72.8 (65.8, 79.1)	65.9 (57.7, 72.5)	74 (66.8, 80.2)	**<.001**
Men (%)		4493 (58)	1837 (74)	933 (53)	420 (48)	442 (54)	596 (47)	**<.001**
BMI (kg/m2)	62	28.1 (24.7, 32.1)	27.2 (24, 30.9)	28.2(24.7, 32.4)	29.6 (26, 33.5)	29.2 (26.1, 33.7)	28.6 (25.1, 32.5)	**<.001**
Underweight (<20) (%)		281 (4)	118 (5)	73 (4)	14 (2)	11 (1)	31 (3)	
Lean (20-24.9) (%)		1723 (23)	677 (27)	379 (22)	135 (16)	132 (16)	267 (22)	
Overweight (25-29.9) (%)		2751 (36)	914 (37)	635 (37)	304 (35)	301 (37)	436 (35)	**<.001**
Obese (30-39.9) (%)		2401 (32)	691(28)	531 (31)	339(39)	304(37)	428 (35)	
Morbidly Obese (>=40) (%)		421 (6)	71 (3)	119 (7)	74 (9)	63 (8)	72 (6)	
NYHA class (%)	189							
I		1936 (26)	346 (14)	362 (21)	311 (37)	378 (50)	425 (35)	
II		3379 (45)	1174 (47)	835 (48)	398 (47)	277(36)	542 (45)	**<0.001**
III		1986 (27)	886 (36)	513 (29)	130 (15)	101 (13)	230 (19)	
IV		149 (2)	75 (3)	31 (2)	6 (1)	5 (1)	13 (1)	
Systolic BP (mmHg)	24	139 (122, 158)	129 (113, 145)	142 (126, 162)	150 (133, 166)	143 (129, 158)	148 (130, 165)	**<0.001**
Oedema	686							
None		4087 (59)	1244 (55)	753 (46)	528 (65)	578 (78)	736 (65)	
Trace		944 (14)	343 (15)	241 (15)	120 (15)	62 (8)	135 (12)	**<0.001**
Ankles		1398 (20)	472 (21)	471 (29)	126 (15)	79 (11)	170 (15)	
Above ankles		524(8)	185(8)	160(10)	41(5)	21(3)	84(7)	
Left ventricular systolic dysfunction								
LV Impairment	1							

None		3194(45)	0(0)	969(56)	626(72)	697(86)	902(72)	
Trivial		560(8)	0(0)	265(15)	94(11)	60(7)	141(11)	
Mild		931(13)	0(0)	512(29)	147(17)	58(7)	214(17)	
Worse		2485(35)	2485(100)	0(0)	0(0)	0(0)	0(0)	
EF Simpsons (%)	3506	47(37,58)	33(26,38)	53(47,60)	57(50,62)	59(54,64)	55(48,61)	
Findings on electrocardiogram								
Heart rate (bpm)	209	72(62,84)	75(64,88)	72(62,85)	67(58,78)	71(62,81)	71(62,82)	**<0.001**
Heart rhythm (Sinus) (%)	190	4976(67)	1517(62)	785(45)	783(91)	778(97)	869(72)	**<0.001**
QRS (ms)	356	98(86,118)	114(98,144)	96(86,112)	92(82,100)	90(82,98)	92(84,114)	**<0.001**
Blood test								
Haemoglobin (g/dL)	1397	13.4(12.2,14.5)	13.5(12.2,14.6)	12.9(11.7,14.1)	13.5(12.6,14.4)	14(13.2,15)	13.3(12.1,14.4)	**<0.001**
Sodium (mmol/L)	1229	139(137,140)	139(136,140)	139(137,140)	139(137,141)	139(138,141)	139(137,140)	**<0.001**
Potassium (mmol/L)	1270	4.3(4,4.6)	4.4(4.1,4.7)	4.3(4,4.7)	4.3(4.1,4.6)	4.3(4,4.5)	4.3(4,4.6)	0.39
Urea (mmol/L)	1229	6.5(5,8.9)	7.2(5.4,10.1)	7.1(5.4,9.8)	5.9(4.7,7.3)	5.2(4.2,60.3)	6.2(4.8,8.2)	**<0.001**
Creatinine (umol/L	1237	97(80,121)	106(88,135)	100(82,129)	87(74,103)	82(71,96)	92(77,114)	**<0.001**
Albumin (g/L)	1331	38(35,40)	38(35,40)	37(35,39)	39(37,41)	40(38,42)	38(35,40)	**<0.001**
NT-proBNP (ng/L)	1116	792(219,2165)	1752(745,3900)	1269(727,2364)	225(165,296)	60(37,91)		
Heart failure medication								
Loop Diuretic (%)		4568(60)	1975(79)	1204(69)	337(39)	209(26)	554(44)	**<0.001**
Furosemide EquivDailyDose (mg)		40(40,80)	40(40,80)	40(40,80)	40(40,40)	40(40,40)	40(40,80)	**<0.001**
Thiazide (%)		561(7)	107(4)	136(8)	104(12)	94(12)	102(8)	**<0.001**
Beta-blocker (%)		4333(57)	1875(75)	1077(62)	407(47)	211(26)	499(40)	**<0.001**
ACE/ARB (%)		5166(68)	2185(88)	1200(69)	501(58)	352(43)	650(52)	**<0.001**
MRA (%)		1584(21)	981(39)	302(17)	65(8)	27(3)	89(7)	**<0.001**

Table 6.1: Figure 6.8: Baseline demographic and clinical characteristics. See Table 4.6 for abbreviations: HeFREF, HF with reserved ejection fraction, HeFPEF, HF with preserved left ventricular ejection fraction (type 1 is defined as echocardiographic abnormalities that could account for symptoms and NT-

proBNP concentration >400 ng/L, and type 2 is defined as no LVSD but NT-proBNP concentration between 400 - =125 ng/L. Bold text indicates statistical significance at the 0.05 level (two-tailed). *NT-proBNP testing was introduced and became a routine part of clinical services while the data for the study was being collected. Initially, this test was not available, but it was implemented during the ongoing data collection period.

6.5.1.2 Patient follow-ups (OPD), admissions and deaths

Every patient had an initial outpatient department (OPD) visit from where the transition started. Out of 7,496 patients, 3,485 had subsequent OPD follow-ups. A total of 11,103 OPD follow-ups were noted over 24 months. On average, there were 3.18 follow-ups per patient. When selecting only the latest OPD follow-up within a period, 8,241 follow-ups were counted, averaging 2.36 per patient. Cardiovascular (CV) hospital admissions were noted for 3,302 patients, with the latest admission per cycle bringing the count to 5,386, averaging 1.63 admissions per individual. The overall admission count was 6,369, with an average of 1.92 per patient. Within the 24-month period, 1,372 deaths occurred, representing 18% of the total patient population**.**

6.5.2 Probabilistic estimation of patients' risk using AMC

An AMC allows for estimating how many cycles (time intervals) a patient will stay in transient states before eventually moving to one of the absorbing states. It also estimates the likelihood and proportions of patients ending up in each of these absorbing states. As stated above, the transition matrix for any absorbing chain to have a canonical form with four block matrices, $I, 0, R$ and Q, as shown in equation (6.9). The partition of transition probability matrix is such that the first rows and columns represent the absorbing states, while the partition Q represents the N non-absorbing (i.e., transient) states.

$$\begin{array}{c} \\ A \\ N \end{array}\begin{array}{c} \begin{array}{cc} A & N \end{array} \\ \left[\begin{array}{c|c} I & 0 \\ \hline R & Q \end{array}\right] \end{array} = \mathcal{P} \tag{6.9}$$

Where, I is an identity matrix, and 0 is $zeros$ materix, R is a non-zero N-by-A matrix (non-absorbing to absorbing states) and Q is a N-by-N matrix (non-absorbing to non-absorbing). The states are ordered such that the absorbing states come first followed by the transient states. The risk states of this study are ordered as $[Dead(D),\ Left\ (L),\ Hosp(H),\ OPD(O),\ No\ event\ (NE)]$ and representing P_{obs}^{1} into a canonical form with four block matrices (P) is as described with example in Table 6.4.

6.5.3 Distribution and proportion at initial two transitions

Table 6.2 presents the distribution and proportion of patients after the first transition from BL to the end of 1st cycle. Initially, all subjects were in the $[OPD]$ state. Over the next four months, they transitioned into one of five possible states with the following proportion: 6% (N =427) had died, 25% (N = 1842) had left the service, 21% (N = 1559) were admitted to hospital, 30% (N = 2254) had attended the OPD clinic again and 19% (N = 1414) did not access the service. N is total number of patients.

1st cycle (N at baseline 7,496)		Absorbing states		Transient states		
		[Dead]	*[Left]*	*[Hosp]*	*[OPD]*	*[No_Event]*
Initial	D/P	427(0.06)	1842(0.25)	1559(0.21)	2254(0.3)	1414(0.19)

Table 6.2: Distribution and probability observed (from data) at 1st transition

Abbreviation D = Distribution and P = Probability

Table 6.3 displays the observed frequencies from 1st cycle to the end of 2nd cycle. Patients in [*Dead*] (n=427) and [*Left*] (n=1842) states at the end of 1st cycle remain in these states at the next follow-up. Of 1,559 hospitalised patients, 104 patients died, 266 left the service, 376 were hospitalised again, 233 were seen in out-patients and 328 did not attend the service.

States		To					Total
		[Dead]	***[Left]***	***[Hosp]***	***[OPD]***	***[No Event]***	
From	***[Dead]***	427					427
	[Left]		1842				1842
	[Hosp]	104	266	376	223	590	1559
	[OPD]	54	188	273	493	1246	2254
	[No Event]	66		269	190	889	1414

Table 6.3: Frequencies observed from the data at 2nd transition

6.5.4 Long-term behaviour

Table 6.4 shows the probabilities of 2nd transition in four block matrices as explained in the equation 6.9. Patients who died or left the system in the 1st cycle remain in those states with probability 1. As examples of transitions between transient states, patients in the [*Hosp*] state had a probability of being [*Dead*] of 0.07; and patients in the [OPD] state had a probability of being [*Hosp*] of 0.12.

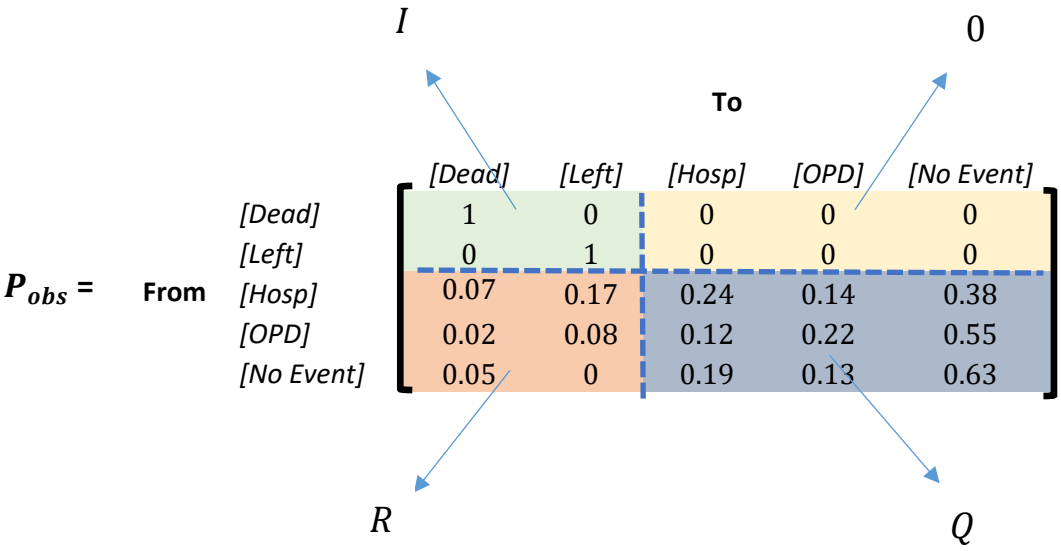


Table 6.4: Transition probabilities during the 2nd cycle

A 5×5 matrix, represents the observed probabilities **(P_{obs})**. The states are ordered such that absorbing states (**ABS**) come first and then the transient states (**TR**). The different colours represent the canonical form of the Absorbing Markov chains with four block ***I, O, R and Q*** matrices.

The corresponding transition probabilities (P_{obs}) seen in Table 6.4 can be represented with four block matrices, *I, 0, R and Q*:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ 0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ R = \begin{bmatrix} 0.07 & 0.17 \\ 0.02 & 0.08 \\ 0.05 & 0 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 0.24 & 0.14 & 0.38 \\ 0.12 & 0.22 & 0.55 \\ 0.19 & 0.13 & 0.63 \end{bmatrix} \quad (6.10)$$

The standard form of (P_{obs}) is very useful in determining the limiting matrix ($\bar{P}$) for an absorbing Markov chain, where P^k approaches a limiting matrix $\bar{P}$ as k increase:

$$\begin{bmatrix} I & 0 \\ FR & Q \end{bmatrix} = \bar{P} \tag{6.11}$$

The rows of both R and FR corresponding to the transient (non-absorbing) states. To determine the sub-matrix FR of limiting matrix $\bar{P}$, the *fundamental matrix F (Yashinski, 2021; Sargent et al., 2024)* has to be determined and is given by:

$$F = (I - Q)^{-1} \tag{6.12}$$

Where I and Q are as before with the appropriate dimensions, thus using the submatrices given above, it can be seen that:

$$F = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.24 & 0.14 & 0.38 \\ 0.12 & 0.22 & 0.55 \\ 0.19 & 0.13 & 0.63 \end{bmatrix} \right)^{-1} \tag{6.13}$$

$$F = \begin{bmatrix} 2.65 & 1.24 & 4.56 \\ 1.82 & 2.55 & 5.66 \\ 2 & 1.53 & 7.03 \end{bmatrix} \tag{6.14}$$

Matrix F represents square matrix where both rows and columns represent the transient states, i.e., $[Hosp, OPD, No\ Event]$. This means that row i corresponds to the ith transient state, not just the ith state in general. Consequently, $F(i, j)$ represent the expected number of cycles that the chain spends in the jth transient state given that the chain started in the ith transient states. The sum of entries in F determined by:

$$F(i,j) = Q^0(i,j) + Q^1(i,j) + Q^2(i,j) + \cdots \quad (6.15)$$

where $Q^t(i,j)$ represent the probability that the process which started in the ith non-absorbing state will be in jth non-absorbing state in period t. $Q^t(i,j)$ can also be explained as the expected $proportion$ of period t spent in the jth state. Summing over all time period t gives average number of cycles that it take to go from a given non-absorbing states to an absorbing state, given that the chain started in the ith non-absorbing state. The equation can also be written as:

$$\%\ expected\ number\ of\ cycles\ before\ absorption = \sum_{j=1}^{3}(I-Q)^{-1} = \begin{matrix} Hosp \\ OPD \\ No\ Event \end{matrix}\begin{bmatrix} 8.45 \\ 10.03 \\ 10.55 \end{bmatrix} \quad (6.16)$$

The F matrix (equation 6.14) provide the expected number of visits to each transient states until it entered in to one of the absorbing state. For instance, the first row shows that if the patient is in the $[Hosp]$ state after their initial transition, they will on average spend approximately 3 cycles in this state, 1 cycle in the $[OPD]$ state and 5 cycles without requiring HF services before reaching an absorbing state. Similarly, the second row indicates that if the patient begin in the $[OPD]$ state after their initial transition, they are expected to spend an average 2, 3 and 6 cycles in the $[Hosp]$, $[OPD]$ and $[No\ Event]$ states, respectively. The F matrix provides an estimate of the number of cycles until a subject enterred an absorbing state, obtained by summing each row of F, shown in equation (6.16).

The limiting nature of AMC is like a cumulative utility where Markov cohort is completely absorbed in non-transient states. The long term behaviour of a Markovian model can be shown by the limiting matrix equation 6.11. Thus the limiting matrix, FR is

$$F \cdot R = \begin{bmatrix} 2.65 & 1.24 & 4.56 \\ 1.82 & 2.55 & 5.66 \\ 2 & 1.53 & 7.03 \end{bmatrix} \cdot \begin{bmatrix} 0.07 & 0.17 \\ 0.02 & 0.08 \\ 0.05 & 0 \end{bmatrix} \tag{6.17}$$

$$FR = \begin{bmatrix} 0.44 & 0.55 \\ 0.46 & 0.51 \\ 0.52 & 0.46 \end{bmatrix} \tag{6.18}$$

By putting FR value into equation (6.11), the Limiting matrix ($\bar{P}$) of this analysis is as shown in equation 6.18.

The limiting distribution for an absorbing chain generally relies on the initial state of the process as shown in the Table 6.2, which refers to the cohort of patients at initial transition. In contrast, the limiting distribution for a regular chain does not depend on the initials state.

$$\bar{P} = \begin{array}{c} \\ [Dead] \\ [Left] \\ [Hosp] \\ [OPD] \\ [NE] \end{array} \begin{array}{c} \begin{array}{ccccc} [Dead] & [Left] & [Hosp] & [OPD] & [NE] \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.44 & 0.55 & 0 & 0 & 0 \\ 0.46 & 0.51 & 0 & 0 & 0 \\ 0.52 & 0.46 & 0 & 0 & 0 \end{bmatrix} \end{array} \tag{6.19}$$

The *limiting matrix* ($\bar{P}$) (equation (6.19)) displays the probabilities of patients reaching either of the two absorbing states, based on the state they are in after the first cycle. To fully understand equations (6.16) and (6.19) requires to interpret together. An illustration of

interpreting the results suggested that the model predicts that a patient who was hospitalised after the first cycle has a 0.44 probability of dying *within (approximately) 8 cycles*. Here 8 cycles equates to further 6 cycles beyond the initial two cycles which were used the derived the model. It is important to note that, according to the model's description, every individual will inevitably reach an absorbing state within the specified timeframe (death or the end of the model.

6.5.5 Short term behaviour

Transition probabilities observed from the data and predicted by the model for all patients up to the 6th cycle (24 months) presented in Table 6.5. it is worth mentioning again that the probabilities predicted by a model solely based on the observed data from initial two transitions. The table also indicates the extent to which the prediction differ from the observed data.

The predicted probabilities of the model shown in the left-hand columns, the centre columns represented the observed probabilities, and the error (E) between two shown in right-hand columns. E.g.,, at the 4th cycle, the model predicted the following, 14% of individuals will be deceased, 37% will have discharged, 10% will be hospitalised, 8% will be attended OPD clinic and 31% will not require any HF service. A heat map of colour coding from green to red indicates the increasing difference. Negative signs show underestimation while positive signs show overestimation. It can be seen from the data for cycle 3 to 6 (after the initial assessment), there is high level of agreement regarding important clinical states, death and hospitalisation.

From		Prediction (model)					Observed (from data)					Error (E) between prediction				
	Cycle	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*
BL	**1**	-	-	-	-	-	0.06	0.25	0.21	0.30	0.19	-	-	-	-	-
	2	-	-	-	-	-	0.09	0.31	0.12	0.12	0.36	-	-	-	-	-
	3	0.12	0.34	0.11	0.09	0.34	0.11	0.34	0.11	0.26	0.18	0.00	0.00	0.00	-0.17	0.16
	4	0.14	0.37	0.10	0.08	0.31	0.14	0.41	0.10	0.15	0.21	0.01	-0.04	0.01	-0.07	0.10
	5	0.16	0.39	0.09	0.07	0.28	0.16	0.47	0.10	0.07	0.20	0.00	-0.08	0.00	0.00	0.07
	6	0.19	0.41	0.08	0.07	0.25	0.18	0.54	0.08	0.19	0.00	0.00	-0.13	0.00	-0.13	0.25

Table 6.5: Predicted vs observed probabilities for overall population

Probabilities are shown (up to 6 cycles) and are rounded to 2 decimal points

6.5.6 Long-term survival analysis

Equations 6.16 and 6.19 shows the long-term prediction until the probabilities of patients transitioning to one of the two absorbing states. The probability of survival over 9 cycles is illustrated in Figures 6.8 and 6.9.

Figure 6.8 is a graph that show the Kaplan-Meier survival curves (Goel et al., 2010) for three groups of patients in transient states ([*Hosp*], [*OPD*] and [*No Event*] at 4 month's interval. The probability at each cycle represents survival probability at that point. For example, the model predicts the survival probability 0.57, 0.55 and 0.50 at cycle 9th, 10th and 11th for patients in ([*Hosp*], [*OPD*] and [*No Event*], respectively.

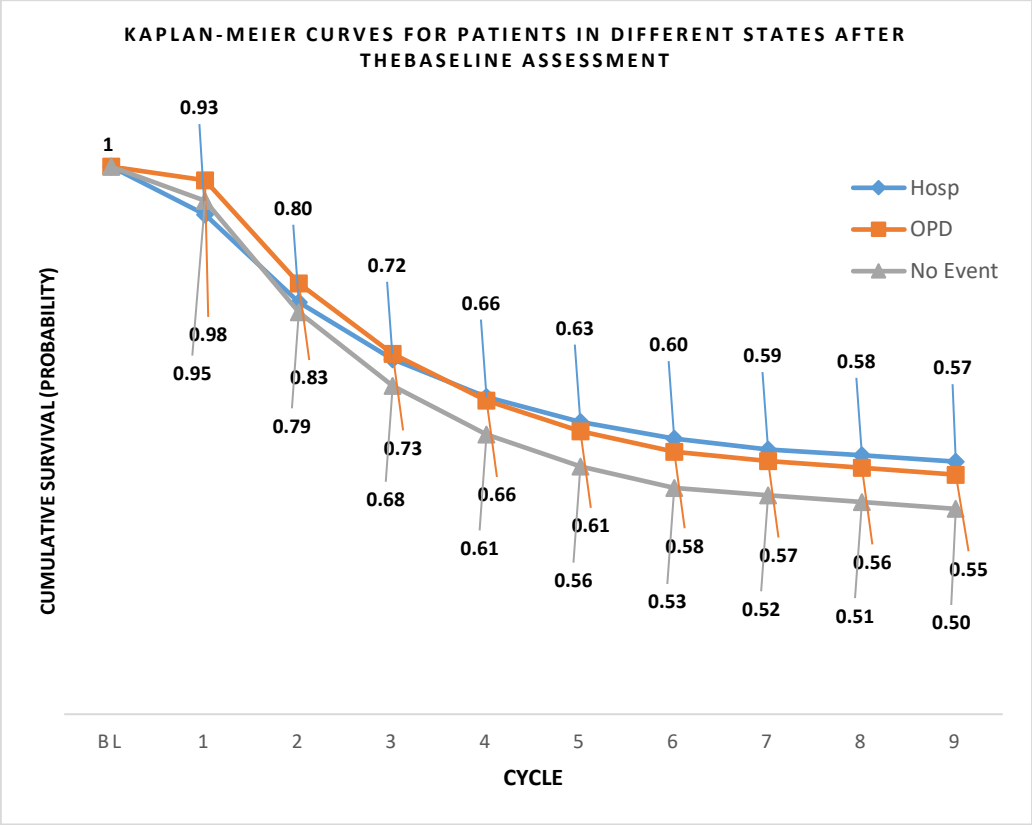


Figure 6.8: Survival curves for three groups of patients in transient states (i.e.,[Hosp], [OPD] and [No Event] at 4 month's interval.

I analysed the observed the probabilities of patients other than controls. As seen in Table 6.6, of the patients who were in the [*Hosp*] state at 4 month's interval, 39% had died within the period of 48 months. Similarly, for those who had an outpatient follow-up at 4 month's interval 32 % were dead within 52 months.

Starting state at 4 months' cycle	Death all cause	Cycle No (months)
[*Hosp*]	39%	12 (48 M)
[*OPD*]	32%	13 (52 M)
[*NE*]	40%	14 (56 M)

Table 6.6: Mortality rate absorbed in HLL based on 4moths' transitionThis includes everyone other than controls. Patients were stratified by their starting states.

Generally, long-term predictions are made based on the treatments that patients receive. The treatment is not applied in the current analysis. In subsequent analyses (chapters 7 and 8), treatments and other covariates will be incorporated into model to examine the long-term effects of treatment. The other reason for this is that as the number of patients in the cycles come down, the errors in the predictions increase, thus the difference between the predictions and the true numbers. These differences can be explained by the presence of control patients in the data.

6.5.7 Prediction based on demographics

AMC was developed for each sub-group for each sub-group to assess differences in the progression of patients by sex and age-groups (those aged 75 and above years old and those under 75 years old)). The model accurately predicted both death and hospitalisation. The fundamental (F) and limiting matrices ($\bar{P}$) showing the long-term prediction for the subgroups (sex and age-groups): for male (m) it is equation 6.20, for female (f) it is equation 6.21, for age-group ≥ 75 it is equation 6.22 and < 75 it is equation 6.23: females spend a lesser number of cycles in the transient states compared to males and have lower likelihood of death. Likewise, patients under 75 years old had a better expected outcome. Table 6.7 – 6.10 displays the observed and predicted probabilities of transition for sub-groups up to the 6^{th} cycle (24 months). E.g.,, Table 6.7 illustrate that at cycle 4, the model predicted the following, 16% individuals deceased, 31% discharged, 11% hospitalised, 10% attending OPD clinic and 31% not requiring heart failure service. A heat map of colour coding from green to red indicates the increasing difference. Negative signs show underestimation while positive signs show overestimation. Tables 6.8 – 6.10 can be interpreted similarly.

$$F_m = \begin{matrix} \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 2.76 & 1.49 & 4.77 \\ 1.92 & 2.8 & 5.76 \\ 2.03 & 1.78 & 6.89 \end{bmatrix} \end{matrix} = \begin{matrix} Total\ cycles \\ \begin{bmatrix} 9.02 \\ 10.48 \\ 10.70 \end{bmatrix} \end{matrix}$$

$$\bar{P}_m = \begin{matrix} \\ [Dead] \\ [Left] \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Dead] & [Left] & [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.50 & 0.50 & 0 & 0 & 0 \\ 0.54 & 0.46 & 0 & 0 & 0 \\ 0.59 & 0.41 & 0 & 0 & 0 \end{bmatrix} \end{matrix} \quad (6.20)$$

$$F_f = \begin{matrix} \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 2.56 & 1.1 & 4.33 \\ 1.75 & 2.45 & 5.58 \\ 2.06 & 1.45 & 7.39 \end{bmatrix} \end{matrix} = \begin{matrix} Total\ cycles \\ \begin{bmatrix} 7.99 \\ 9.78 \\ 10.90 \end{bmatrix} \end{matrix}$$

$$\bar{P}_f = \begin{matrix} \\ [Dead] \\ [Left] \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Dead] & [Left] & [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.33 & 0.67 & 0 & 0 & 0 \\ 0.34 & 0.66 & 0 & 0 & 0 \\ 0.39 & 0.61 & 0 & 0 & 0 \end{bmatrix} \end{matrix} \quad (6.21)$$

$$F_{\geq 75} = \begin{matrix} \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 2.57 & 1.21 & 4.49 \\ 1.75 & 2.54 & 5.86 \\ 1.92 & 1.52 & 7.07 \end{bmatrix} \end{matrix} = \begin{matrix} Total\ cycles \\ \begin{bmatrix} 8.27 \\ 10.15 \\ 10.51 \end{bmatrix} \end{matrix}$$

$$\bar{P}_{\geq 75} = \begin{matrix} \\ [Dead] \\ [Left] \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Dead] & [Left] & [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.47 & 0.53 & 0 & 0 & 0 \\ 0.50 & 0.50 & 0 & 0 & 0 \\ 0.56 & 0.44 & 0 & 0 & 0 \end{bmatrix} \end{matrix} \quad (6.22)$$

$$F_{<75} = \begin{matrix} \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 3.35 & 1.94 & 5.87 \\ 2.53 & 3.25 & 6.83 \\ 2.85 & 2.39 & 8.42 \end{bmatrix} \end{matrix} = \begin{matrix} Total\ cycles \\ \begin{bmatrix} 11.16 \\ 12.61 \\ 13.66 \end{bmatrix} \end{matrix}$$

$$\bar{P}_{<75} = \begin{matrix} \\ [Dead] \\ [Left] \\ [Hosp] \\ [OPD] \\ [NE] \end{matrix} \begin{matrix} \begin{matrix} [Dead] & [Left] & [Hosp] & [OPD] & [NE] \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.22 & 0.78 & 0 & 0 & 0 \\ 0.25 & 0.75 & 0 & 0 & 0 \\ 0.28 & 0.72 & 0 & 0 & 0 \end{bmatrix} \end{matrix} \quad (6.23)$$

From	To	Prediction					Observed					Error				
	Cycle	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*
BL	1	-	-	-	-	-	0.06	0.20	0.22	0.34	0.18	-	-	-	-	-
	2	-	-	-	-	-	0.09	0.26	0.13	0.13	0.39	-	-	-	-	-
	3	0.13	0.28	0.12	0.11	0.36	0.12	0.29	0.11	0.30	0.18	0.01	0.00	0.01	-0.20	0.18
	4	0.16	0.31	0.11	0.10	0.32	0.15	0.35	0.10	0.18	0.22	0.01	-0.04	0.01	-0.09	0.10
	5	0.19	0.33	0.10	0.09	0.29	0.17	0.41	0.11	0.08	0.23	0.02	-0.08	0.00	0.00	0.07
	6	0.22	0.35	0.09	0.08	0.26	0.19	0.49	0.09	0.23	0.00	0.02	-0.13	0.00	-0.15	0.26

Table 6.7: Predicted vs observed probabilities for **male** population (up to 6 cycles)

From	To	Prediction					Observed					Error				
	Cycle	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*
BL	1	-	-	-	-	-	0.05	0.31	0.19	0.24	0.21	-	-	-	-	-
	2	-	-	-	-	-	0.08	0.38	0.11	0.11	0.33	-	-	-	-	-
	3	0.09	0.42	0.10	0.08	0.31	0.11	0.41	0.11	0.19	0.17	-0.01	0.00	-0.01	-0.12	0.14
	4	0.11	0.45	0.09	0.07	0.28	0.12	0.49	0.09	0.11	0.19	-0.01	-0.04	0.00	-0.04	0.10
	5	0.13	0.47	0.08	0.06	0.26	0.15	0.55	0.08	0.05	0.17	-0.02	-0.07	0.00	0.01	0.09
	6	0.14	0.50	0.08	0.05	0.23	0.17	0.61	0.07	0.14	0.00	-0.02	-0.12	0.00	-0.09	0.23

Table 6.8: Predicted vs observed probabilities for **female** population (up to 6 cycles)

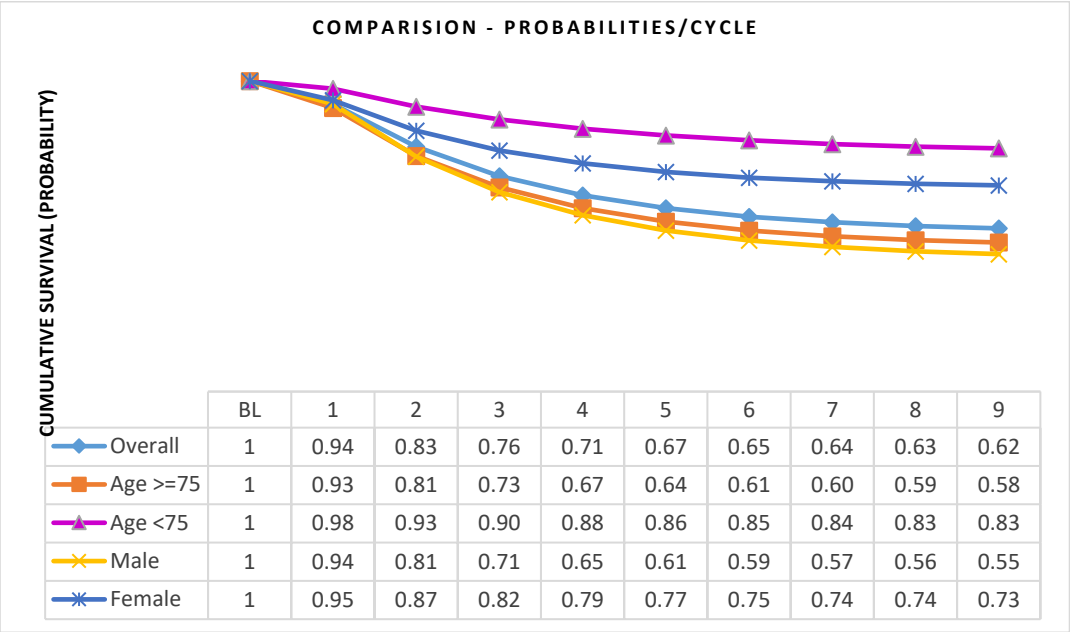
From	To	Prediction					Observed					Error				
	Cycle	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*
BL	1	-	-	-	-	-	0.07	0.22	0.21	0.30	0.20	-	-	-	-	-
	2	-	--	-	-	-	0.10	0.28	0.12	0.12	0.37	-	-	-	-	-
	3	0.14	0.31	0.11	0.09	0.35	0.13	0.32	0.11	0.25	0.18	0.00	-0.01	0.00	-0.16	0.17
	4	0.17	0.34	0.10	0.08	0.32	0.16	0.39	0.10	0.15	0.21	0.01	-0.05	0.00	-0.07	0.11
	5	0.19	0.36	0.09	0.07	0.28	0.19	0.45	0.10	0.07	0.20	0.00	-0.09	-0.01	0.01	0.09
	6	0.22	0.38	0.08	0.07	0.26	0.22	0.52	0.08	0.18	0.00	0.00	-0.14	0.00	-0.12	0.26

Table 6.9: Predicted vs observed probabilities for **≥ 75** (y) population (up to 6 cycles)

From	To	Prediction					Observed					Error				
	Cycle	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*	*[Dead]*	*[Left]*	*[Hosp]*	*[OPD]*	*[NE]*
BL	1	-	-	-	-	-	0.02	0.31	0.20	0.32	0.15	-	-	-	-	-
	2	-	-	-	-	-	0.04	0.38	0.12	0.12	0.34	-	-	-	-	-
	3	0.05	0.41	0.12	0.10	0.32	0.05	0.41	0.11	0.28	0.15	0.00	-0.01	0.02	-0.18	0.17
	4	0.06	0.44	0.11	0.10	0.30	0.06	0.47	0.09	0.18	0.20	0.00	-0.03	0.02	-0.08	0.09
	5	0.07	0.47	0.10	0.09	0.27	0.07	0.54	0.09	0.08	0.22	0.00	-0.07	0.02	0.00	0.05
	6	0.08	0.50	0.10	0.08	0.25	0.08	0.61	0.08	0.24	0.00	0.00	-0.11	0.01	-0.15	0.25

Table 6.10: Predicted vs observed probabilities for **> 75** (y) population (up to 6 cycles)

Figure 6.9 Kaplan-Meier curves shown below is the probability of survival over 9 cycles. This displays overall survival rates as well as survival rates within different sub-groups without stratifying based on transition at 4 months' cycle. Notably, younger patients and female consistently demonstrate higher survival rates across cycles. This confirms the findings of the previous results.



	BL	1	2	3	4	5	6	7	8	9
Overall	1	0.94	0.83	0.76	0.71	0.67	0.65	0.64	0.63	0.62
Age >=75	1	0.93	0.81	0.73	0.67	0.64	0.61	0.60	0.59	0.58
Age <75	1	0.98	0.93	0.90	0.88	0.86	0.85	0.84	0.83	0.83
Male	1	0.94	0.81	0.71	0.65	0.61	0.59	0.57	0.56	0.55
Female	1	0.95	0.87	0.82	0.79	0.77	0.75	0.74	0.74	0.73

Figure 6.9 show the Kaplan-Meier curves for overall patients and different subgroups.

The Figures 6.10 – 6.13 offer a longitudinal perspective, giving insights into the relative severity of each state with respect to survival. Graphs shown for male, female, patients aged ≥ 65 and patient aged <65, stratified further based on the starting transitional states at 4 months. The trend indicates that patients hospitalised at 4 month cycle have a better survival probability than those who have been to OPD visits or had no events. Similarly, as shown above the young and women had better survival rates across all cycles.

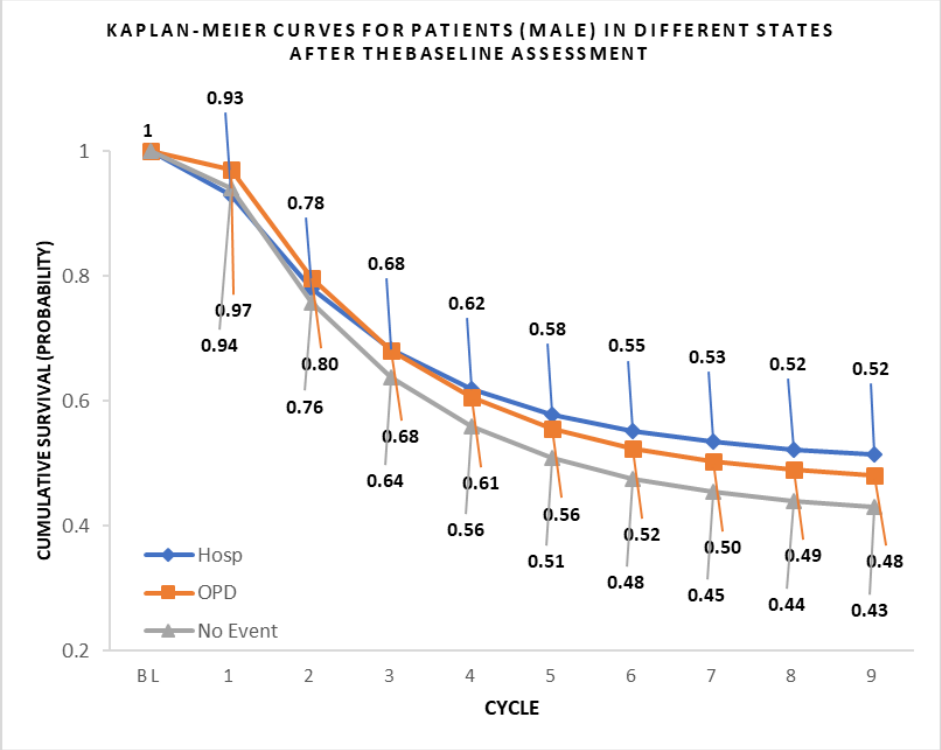


Figure 6.9: Survival curves for **male** population based on their initial transient states

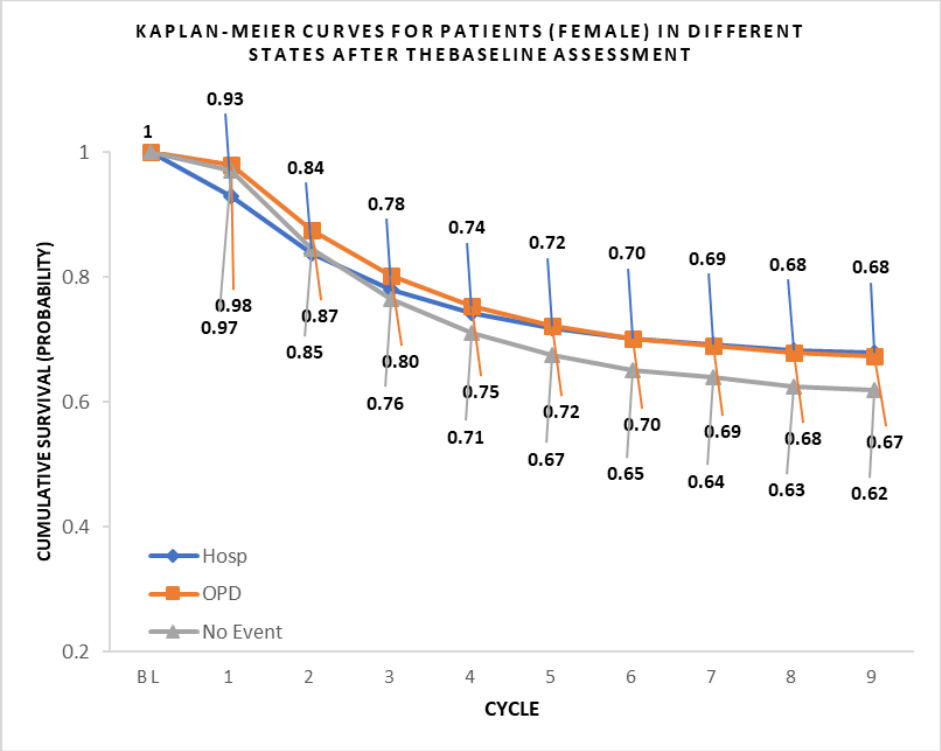


Figure 6.10: Survival curves for **female** population based on their initial transient states

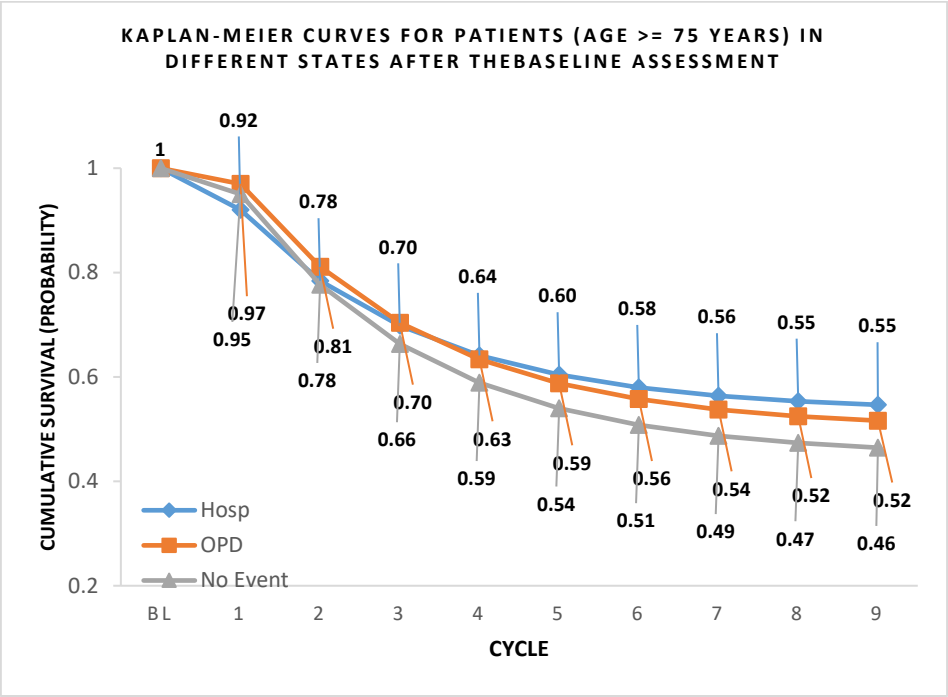


Figure 6.11: Survival curves for **≥ 65 (y)** population based on their initial transient states

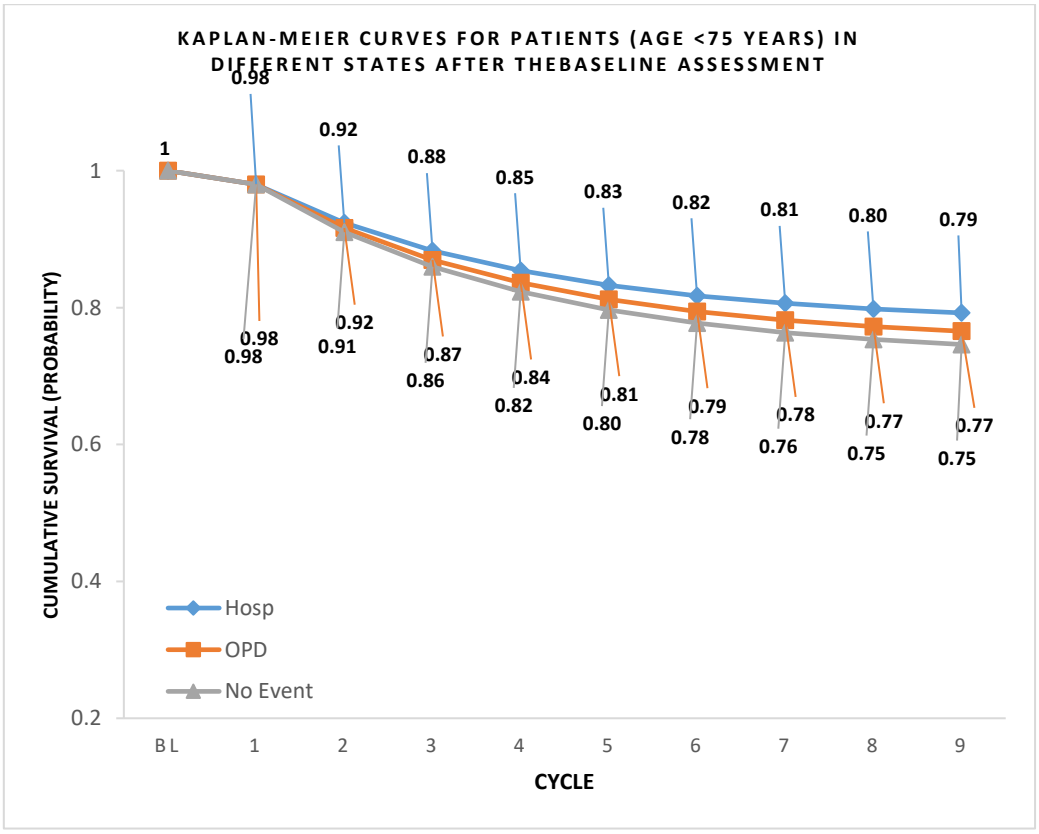


Figure 6.12: Survival curves for **< 75 (y)** population based on their initial transient states

6.6 **Trajectory analysis**

In the context of trajectory analysis the progression of HF disease is shown using the phylogenetic tree structure in this section. Figure 6.13 summarises the course of HF disease in HLL patients across five cycles. This model build on the first-order Markov framework, which groups patients based on their current state without considering the different paths they took to get there. From this state the model then examine their next transition. Same 5 distinct states are used which were described for AMC model (described in section 6.4.2). The purpose and utility of trajectory analysis is provided in chapter 3, section 3.5. Implicit treatment response and any other state-dependence is reflected in the rate of changes.

It can be observed that the transition probabilities to the [*Dead*] state from transient states are somewhat consistent and overlapping. From the [*Hosp*] state, the probability of transitioning to [*Dead*] ranges from 0.05 to 0.07 across all four cycles. Transitions from [*OPD*] to [*Dead*] ranges between 0.02 to 0.03 and from [*NE*] to [*Dead*] they vary from 0.03 to 0.05.

When examining transitions to the [*Hosp*] states the probabilities differ slightly. The probability of readmission (meaning going from [*Hosp*] to [*Hosp*]) falls between 0.21 to 0.26, which is higher than transitioning from [*OPD*] to [*Hosp*] (0.09 to 0.30) and from [*NE*] to [*Hosp*] (0.15 to 0.26). This indicates that patients who are hospitalized are more likely to be readmitted than those transitioning from outpatient from a state where no event occurred. These figures highlight a clear trend of higher readmission probabilities for patients already in the hospital compared to other transitions.

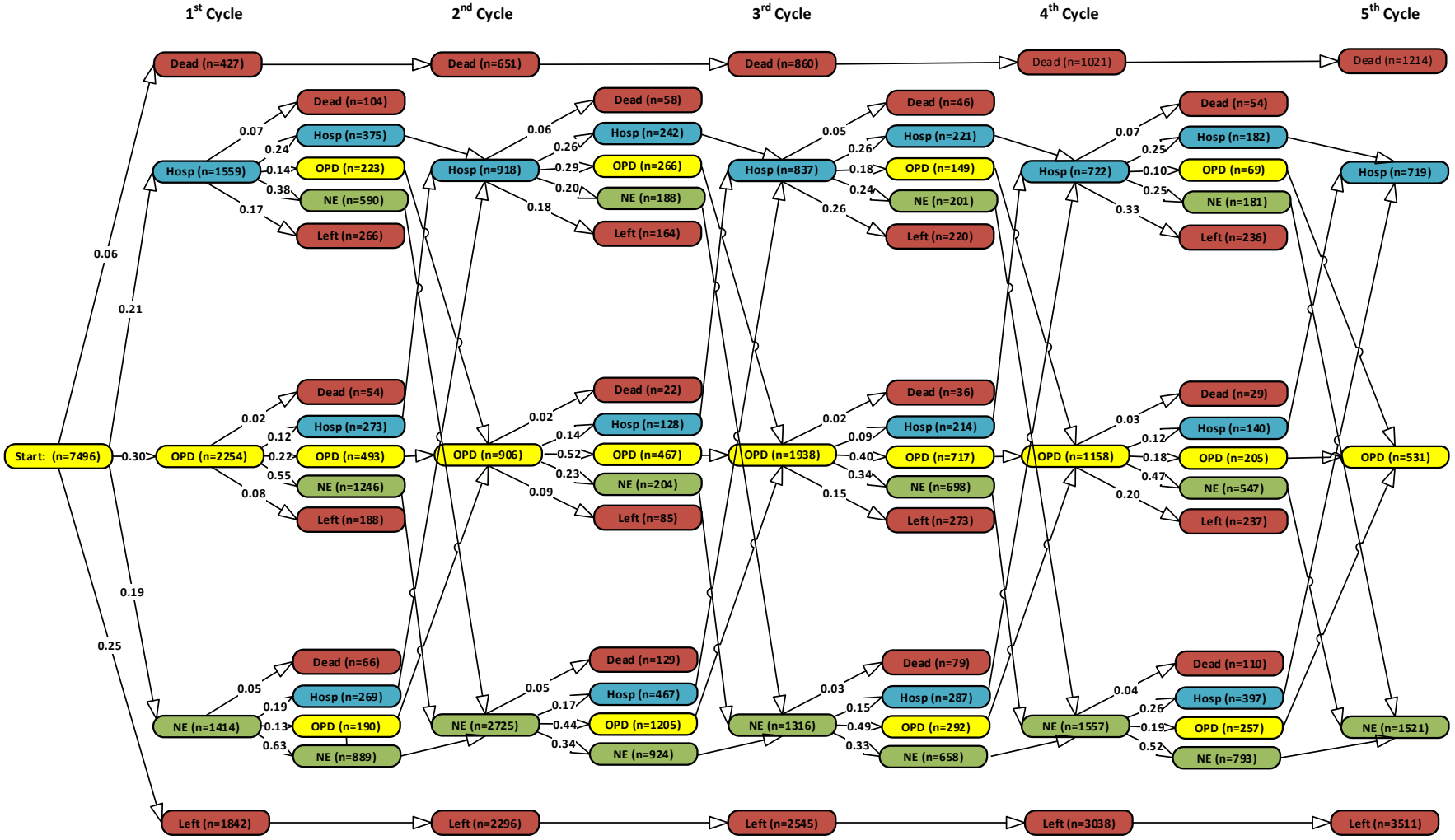


Figure 6.13: Progression of CHF among HLL patients using phylogenetic tree structure. This visual representation helps to trace the various pathways of disease progression within the patient population.

6.7 **Discussion**

In this section I am going to summarise this chapter that what I have done so far and looked at dynamic risk modelling and whole population in general. I have developed models that illustrate both short-term and long-term progression. The survival graphs, along with fundamental and limiting matrices, display the respective percentages, providing quantitative insights into patient outcomes. Additionally, the phylogenetic tree diagrams depict the transitions within the HLL patient population, visually representing the various pathways and states that patients experience over time. These combined approaches offer a comprehensive view of disease dynamics and patient trajectories. My dynamic risk models have shown high level of accuracy for critical states, such as death and hospitalisation. The strong agreement predicted and observed outcomes highlights the models' reliability and effectiveness in forecasting important health events in patients.

This method predicts individual survival and also treats the clinical trajectory of patients as a group. For example current palliative performance scales (PPS) assess the risk of death for individual patients within a specific timeframe, which is useful for those patients. However, they fail to capture disease behaviour patterns at the population level.

The management of CHF disease typically prioritise high risk patients, because targeting intervention at those who are most at risk can help to reduce the number of hospitalisations and mortality. However, it is unclear how effective these management programs over time for patients who are considered to be at lower risk (Chan et al., 2008).

To achieve better long-term results in healthcare management, it is important to address the needs of CHF patient at varying levels of risk by monitoring their risk trajectories and health events over time. This approach allows for a more nuanced and dynamic management of CHF, ensuring that even lower-risk patients receive appropriate care as their condition evolves. My finding align with (Krajewska et al., 2017) and (Zhang et al., 2018a), both stress the importance of regular reassessing patients' risk levels because these risks can change over time. Models developed in this study provided dual temporal perspective on both short-term and long-term predictions underscores it comprehensive utility in patients' health management.

Dynamic risk stratification using AMC to broad group of patients (without selecting specific subgroups) who visited a community HF clinic is complex and ambitious due to several factors. This approach is technically complex because it requires precisely categorising patients' disease progression into a finite number of distinct and comprehensive states. Each state must be mutually exclusive (no overlap between states) and exhaustive (covering all possible conditions that patients might experience). Ambitious because it aims to address two major changelings. The lack of a definitive gold standard for diagnosing HF. The inability of current HF scoring systems to describe disease progression patterns at the population level.

One of the most significant findings of this study was that the event which occurred within the first two cycles (up to 8 months) provided sufficient information to develop a model that could predict events over time with great accuracy, particularly effective in forecasting key clinical outcomes (hospitalisation and mortality).

My Markov model offers a simpler and less computationally intensive method for estimating the probabilities of transitioning between different health states compared to complex scoring models. However, like all data-driven approaches, it requires careful data pre-

processing. This involves understanding the expected clinical progression of the patients and defining the states which clearly represent events during the course of disease. Proper data preparation is crucial for the model to produce meaningful results, which can help improve the management of patients. The study shows that the pattern observed in the first two transition can be reliably applied to predict future outcomes even the risk continuously evolving.

From the result it can be seen that the level of risk can vary depending on the patient's current condition or state, the impact of being in that specific state remains constant. Though the system has no memory – but the present state hold all the information of preceding cycles to predict the future state. I have highlighted the potential advantages of a model that can predict complex problems with minimal computational effort. This model could be beneficial not only for forecasting patient risk levels but also for effectively allocating resources.

I examined patients spanning the entire range of risk, making my study more epidemiologically representative compared to many multicentre studies that enrol patients selectively and not longitudinal. The next step is that I need to check if the same model can be applied to other datasets without additional training. If successful, the model could be further developed to provide personalized predictions.

6.7.1 Limitations

The data for this study was sourced from a single centre, specifically involving patients who were referred for potential HF assessment. While my results indicate that AMC modelling is effective for patients at HLL, it remains uncertain whether these findings can be generalised to other populations or settings. Nevertheless, applying the same methodology to different HF

populations can help evaluate its broader applicability. There might be errors in how the data was coded. Despite this, the predictive model have shown very few errors. The study analysed subgroups based on age and sex, and did not include other clinical variables, such as NT-proBNP. In the following chapter, I incorporate additional variables in the model to improve its accuracy and also enhance its ability to predict events at an individual level.

6.8 **Conclusion**

My finding, that early events in follow up (FU) in heart failure (HF) can strongly predict subsequent outcomes has significant implications for understanding HF progression. HF is typically viewed as a disease with a steady decline, unpredictable hospitalisations, and a constant risk of sudden death (McIlvennan & Allen, 2016). However, the findings indicate that the progression of HF is actually more linear and predictable than previously thought. Understanding this linearity in heart failure progression can lead to improved patient care and outcomes.

Chapter 7 Disease progression in Chronic HF with multistate models

Building on the information provided in previous analysis where I described CHF disease progression using Markov chain model, this chapter employs *Multistate models* (MSM) to model longitudinal data with dependency between observations. Here, I include clinical covariates (such as demography, aetiology, vital signs, blood test results and treatment) to examine their influence on transitions between mutually exclusive clinical states. This allows us to derive the risk of transition (also referred as *hazard for transition)* between states, which in turn allows a greater understanding of the time course of disease progression. In turn, such information might lead to better patient management. These models are efficient because they assume that the state at some arbitrary future time is dependent upon the state during previous time intervals (Castaneda & Gerritse, 2010; Upshaw et al., 2016), thus simplifying statistical analysis (Ma et al., 2015).

I used the '*msm*' package for modelling. The tool operates under the assumption that there is continuous underlying process in the data. I designed the model to reflect this by considering heart failure (HF) disease process as continuous, while clinical events are captured at discrete intervals. I tailored this model to the Markovian structure detailed in chapters 4 and 6.

7.1 Multistate Markov models

A MSM provides a convenient way of modelling prognosis for clinical problems with ongoing risk. In the study of CHF, the changes in a patient's health condition can be described through a finite number of distinct states (i.e. $S = \{s_1, s_2, s_3, \dots, s_n\}$) and maximum of n^2

transition between them. Each state must be mutually *exclusive* (no overlap between states) and *exhaustive* (covering all possible conditions that patients might experience). The factor "*time*" (t) is explicitly associated with the probability of a patient being in certain states (S) in over a sequence of discrete time intervals (are called "cycles"). The underlying disease process is continuous, and clinical events are represented at discrete time points (Gruger et al., 1991; Jackson, 2019). These models are often based on Markovian assumption (i.e., first-order Markov processes). It claims that future progression of the disease process is depends only on the current state (known as "memoryless" property). In other words, the history of the process is summarised by the state occupied at *time* (t). For more detail see chapter 6.

7.1.1 Basic frame for multi-state model

In multi-state models (MSM), the fundamental principle is simple: a subject transitioning *out* of one state must be transitioning *into* another. Unlike in classic Competing risks (CR) models where the subjects are often dropped after the initial transition. Classic Competing risks (CR) models are usually extended form of standard Cox survival models. In MSM when a subject transitions to a new state, the analysis focuses on identifying and understanding the possible subsequent transitions the subject may encounter from this new state onward. This involves considering the risks associated with moving from the current state to other potential states in the model (Sutradhar et al., 2011; Sutradhar & Barbera, 2014; Upshaw et al., 2016; Le-Rademacher et al., 2022). Multi-state modelling framework has extended the desirable qualities of standard Cox regression models or other expansions and enhancements. These models has introduced three innovations to researchers' use of survival models: *transition-specific baseline hazards (also known as transition intensities)*, *transition-*

specific covariates effects, and *transition probabilities*. These innovations collectively provide flexibility and variability to adopt to unfolding of causally complex process over time. To define the basis framework, the following two assumptions have been made about the dependence of the transition rates:

1. Time homogeneous: the intensities remain constant and don't change over time, they are unaffected by the passage of time (t).
2. Markov model: the transition intensities only depend on the current state and the history of the process encapsulated within the current state.

7.1.2 Transition-specific baseline hazards

In a multi-state modelling approach, researchers can adjust or categorize the BL hazard rate for each transition in the model, facilitating a more detailed examination of the data. For each potential transition q, a distinct baseline hazard rate (represented as $\alpha\_\alpha_0(t)$) is calculated, with the parameter t refer to the duration of cycle. This stratification enables the baseline hazard estimates to adapt to complex event sequences, allowing for variations in the estimation process. Quite apposite to standard Cox models, in which only one baseline hazard, denoted as $\alpha_0(t)$; is estimated, without specific distinctions for different transitions q (subscripts) or events.

As depicted in Figure (7.3), the multi-state model permitted unique baseline hazard to each transition. This capability facilitated differentiation among the diverse types of events observed in the data. For instance, my model permitted the rate at which one type of event happens (like transitioning from state 1 to state 2) to differ from the rate at which a different

type of event occurs (such as transitioning from state 1 to state 3), as well as from the rate of recurrence of the same type of event (remaining in state 2). The mathematical formulation (described in later section) enable us to conduct an analysis based on of transition-specific baseline hazards (also known as *transition intensities*) and *transition probabilities*.

7.1.3 Transition intensities of HF models

For the purpose of this thesis, I performed statistical analysis of the data using an extended illness-death model. I described the model as stochastic process$(S(t), 1 \leq t \leq \tau)$, where the value of process at (cycle) time t denotes the state being occupied during that cycle and τ denotes the end of study time (as describe (Putter et al., 2007)). The process is characterized by paths that are continuous from the right and operates within a finite set of states, defined as $S = \{1, 2, 3, 4\}$, representing $[OPD], [Neither], [Hosp]$ and $[Dead]$ states respectively. The transition from one state to another, as well as the timing of these transitions, is determined by a set of *transition intensities* denoted as $q_{j,k}(t, z(t))$ where *j* and *k* represent pair of states. In other words, the *intensities* represents the instantaneous risk (incidence rate) of moving of patients from state *j* to state *k*:

$$q_{j,k}\big(t, z(t)\big) = \lim_{\Delta t \to 0} \frac{P\{S(t + \Delta t) = k | S(t) = j\}}{\Delta t}, j \neq k, \quad (7.1)$$

$$where\ 1\ or\ 0 \leq t \leq \tau\ \ j, k\ \epsilon\ S),$$

This implies that the intensities might rely on either the time of entry (t) into the process or, more broadly, on a range of individual-specific or time-varying explanatory factors

represented as $z(t)$. $S(t)$ is the state occupied at time t. The model of study is based on first-order Markov processes. That is, the state occupied at time $t + \Delta t$ is conditional on the state occupied at time t. The transition intensity (q) represent the entry situated at the intersection of row *j* and column *k* within the transition intensity matrix Q (as indicated in equation 3). The rows of this matrix, the sum up to 0, and conventionally, the diagonal entries of Q matrix are defined as:

$$q_{j,j}(t) = -\sum_{j \neq k} q_{j,k}(t) \tag{7.2}$$

Where $q_{j,k}$ = 0 if a transition from state j to state k is not allowed. My model was developed on Markov assumption, claiming that future evaluation only depends on the current state. This shows that, $q_{j,k}(t, z(t), F_t)$ is not influenced by the observation history F_t of the process up to time just before *t*.

For the model illustrated in Figure 7.3, the intensities obtained (using the equations described above) are summarised in a transition intensity matrix (Q) as:

$$Q = \begin{bmatrix} -(q_{1,2} + q_{1,3} + q_{1,4}) & q_{1,2} & q_{1,3} & q_{1,4} \\ 0 & -(q_{2,3} + q_{2,4}) & q_{2,3} & q_{2,4} \\ 0 & q_{3,2} & -(q_{3,2} + q_{3,4}) & q_{3,4} \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{7.3}$$

Those specific instantaneous state-to-state transitions, which are not permitted in the underlying multi-state model, then the corresponding transition intensities, has a value of 0. The intensities are constant over time when model is time homogeneous, that is, independent of time *t*. The structure defined above allowed us to estimate specific coefficient impacts for each transition identified in the model. It also allowed us to examine whether the influence of

the same covariate differs across various states within a broader process. In the time homogenous model, I have $q_{j,k}(t) = q_{j,k}$. Covariate of interest can be incorporated into the transition intensities using Cox proportional hazards regression model, using the following formation.

$$q_{j,k}(Z) = q_{(j,k),0}\, exp\left(\beta_{j,k}^{T} Z\right) = \exp\left(\beta_{(j,k),0}^{T} + \beta_{j,k}^{T} Z\right) \tag{7.4}$$

Here, $q_{(j,k),0} = \exp\left(\beta_{(j,k),0}^{T}\right.$ is called baseline intensity from state j to state k. The first-order Markov assumption suggests that the future development of the disease is solely determined by its present state, without regard to its past history. This characteristic is referred to as "memoryless" because future changes are independent of previous occurrences, provided the current state is known.

7.2 Model structure and model specification

Patients referred to the HLL out-patient clinic between January 2001 and August 2015 were enrolled in this study. Patients are followed up at regular intervals. Detailed information about the HLL population and out-patient service was provided in the chapter 4 section 4.4. In the previous chapter I developed a dynamic risk model with five distinct states to illustrate both short-term and long-term progression in patients with HF. Patient states were determined at consecutive 4-monthly intervals (cycles) and were followed-up for 24 months. Here my focus shifts to an annual cycle and the follow-up period is extended to five years. This expansion not only broadens the scope of prediction covering both short and long-term behaviour of the

model but also expended model validation through yearly cycles over a span of five years. For this enhanced model, I are considering four risk states and merging the previously separate [*Left*] and [*NE*] states into single risk state.

7.2.1 Diagnostic categories and definitions

For diagnostic categorisation, I followed the same criteria and definition outlined in chapter 6, section 6.4.1. classifying the population into:

1. HeFREF – those with LVEF ≤ 40%.
2. HeFPEF – those with LVEF > 40% and NT-proBNP level:
    a. ≥400 ng/L
    b. 125-399 ng/L
3. No NT-proBNP (only for the cohort with incomplete data) – the diagnosis of HF in this group of patients was uncertain (i.e., those with LVEF > 40% and no NT-proBNP).

Patients with LVEF > 40% & NT-proBNP <125 (ng/L) (referred to as “Controls”), and those with no LVEF information available were excluded as shown in Figure 7.1. This exclusion helped to focus the study to more specific cases of heart failure.

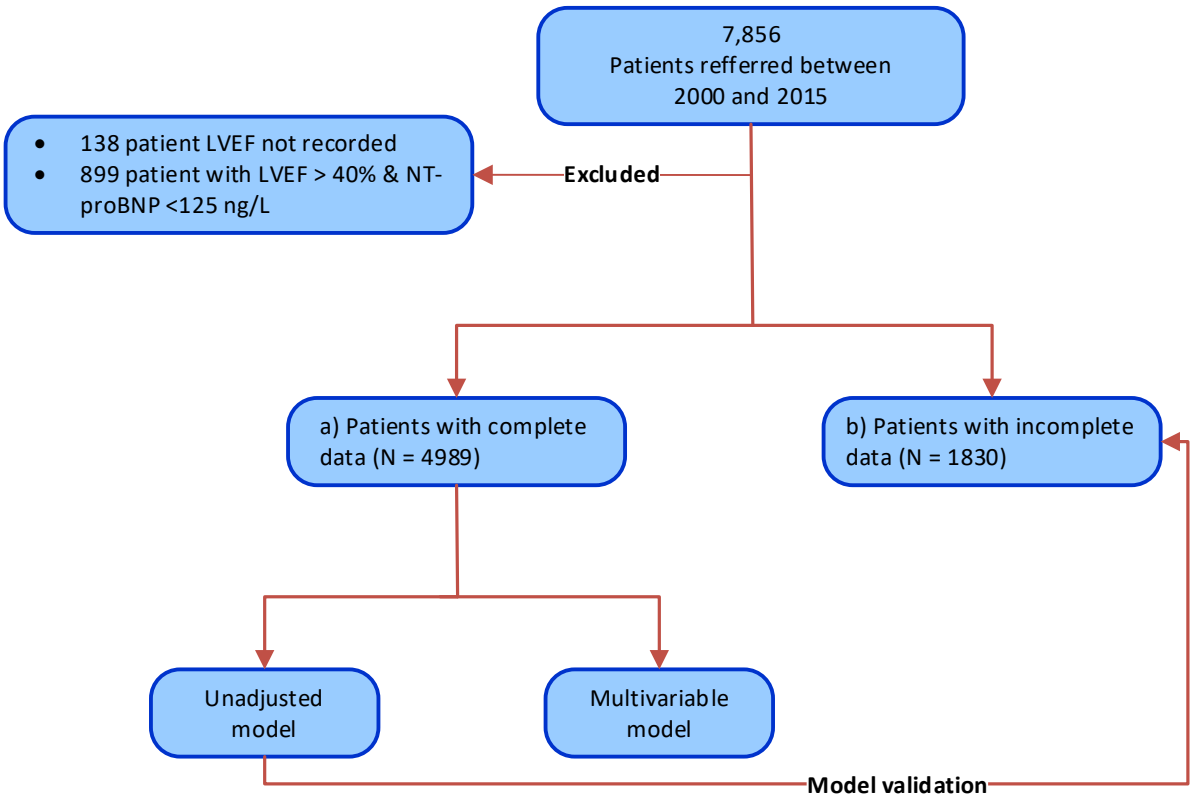


Figure 7.1: Consort diagram illustrated the flow of patients with HFassessed between 2000 and 2015. 1037 patients whose LVEF or LVSD were not available at baseline and patients with LVEF > 40% and Patients with LVEF > 40% & NT-proBNP <125 (ng/L) have been excluded.

Patients who met the study's inclusion criteria are grouped into two cohorts: (a) the cohort with complete cases (i.e., NT-proBNP, haemoglobin, sodium, potassium, urea, creatinine and albumin) b) cohort of patients with incomplete data (where any of the above mentioned variables were missing). The initial model was developed on cohort (a) and I then validated the model on cohort (b). The demographics of these two cohorts are presented in section 8.1. Pictorial representation of study is given in Figure 7.2.

7.2.2 Selection of covariates

I selected common variables to assess their effect on the probability of transitions from one clinical state to another, along with demographic and baseline characteristics (listed in Table 8.1). I chose readily available clinical features known to be related to outcome (Pocock et al., 2013; Nikolaidou et al., 2018; Sokoreli et al., 2018; Koulaouzidis et al., 2019).

7.2.3 Risk states and definition

The patients when first seen for assessment were defined as being in the baseline state [*BL*]. Thereafter, patients could transition into one of the following possible states (*S*) at annual intervals (referred to as successive "cycles"):

- [*Hosp*] – any heart failure hospitalisation during that 1 year cycle;
- [*Dead*] – death (all-cause) during that one year cycle;
- [*Neither*] – patients with no event (heart failure hospitalisation or death) during that one year period.

The model allows the transition to [*Dead*] to occur at any time within an annual cycle. My approach was hierarchical: if a patient was both admitted and died within a single cycle, only the death is considered in the model. The underlying disease process is continuous, and clinical events are represented at discrete time points as described by (Gruger et al., 1991; Jackson, 2019).

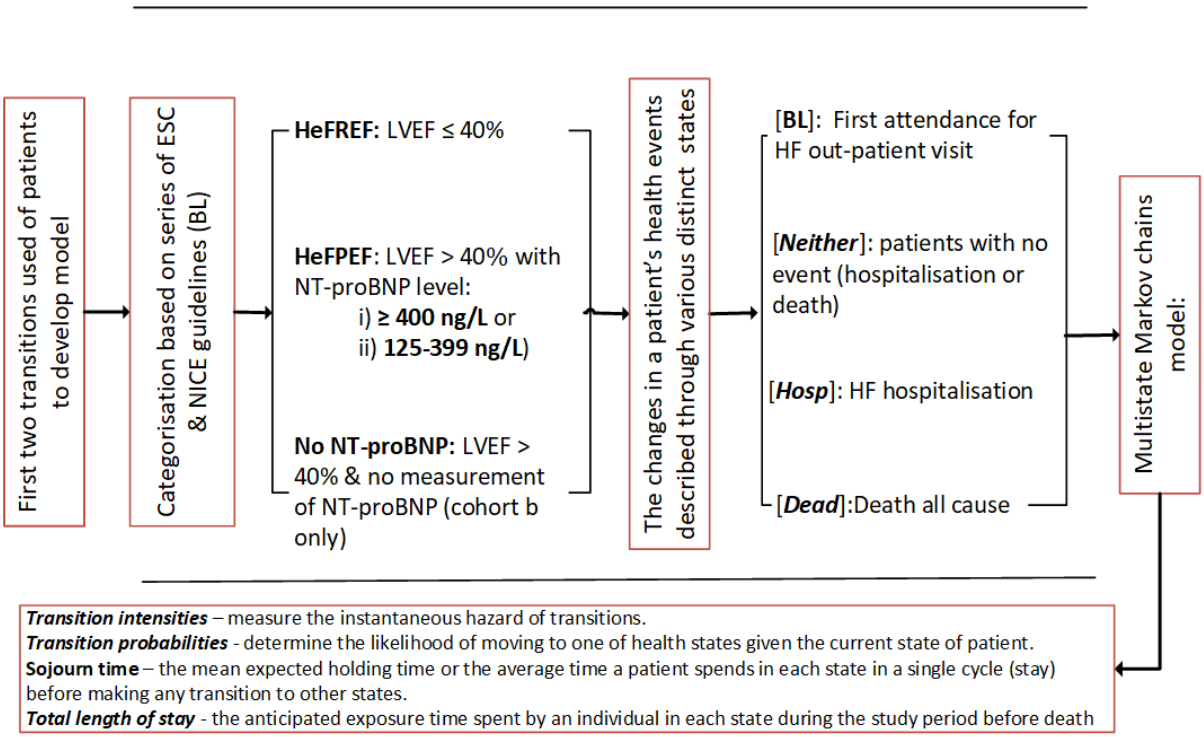


Figure 7.2: Study flow chart for the MSM analyse

.

7.2.4 Disease-driven observation process:

Figure 7.3 shows the four-state model used for describing disease progression. It is an example of a bi-directional non-progressive illness-death model (in other words, patients can transition repeatedly between [*Neither*] and [*Hosp*], but can only transition to [*Dead*] once). The model is based on a first-order Markov process where the state at some arbitrary future time is dependent only upon the immediately preceding state.

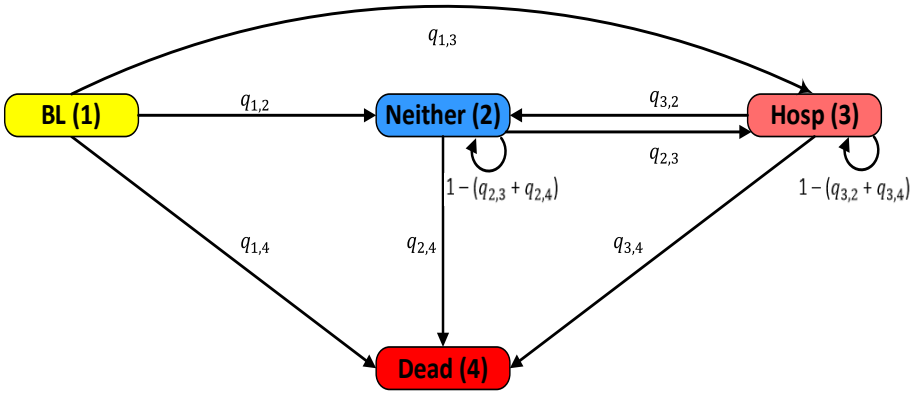


Figure 7.3: A 4 state multistate model.

The arrows in the model indicate the directions in which instantaneous transitions are permitted. From the baseline state [*BL*] transition was unidirectional to any of the three other states. Transitions between [*Neither*] and *[Hosp]* were potentially bidirectional. Once [*Dead*] has been reached, no further transitions were possible (death is an "absorbing state"). The transition intensities (q) are denoted by $q_{j,k}$. State *j (= 1, 2 or 3)* to state *k (= 2, 3 or 4).* [*BL*], [*Neither*], [*Hosp*] and [*Dead*] states are also represented by 1, 2, 3 and 4 respectively. Remaining in the same sate in the next cycle is represented by a minus sign (e.g., $1-(q_{1,2}+q_{1,3}+q_{1,4})$). The transition intensity for remaining in the same state is obtained by subtracting the total of all other probabilities in the same row from 1. This model allows for the analysis of transitions in a simplified yet dynamic manner, capturing the essential features of patient progression through different health states.

7.2.5 Model construction

To allow a multistate model to be fitted to data, figure 7.2 was translated into a square matrix (*Q matrix*) using equation 7.3, taking a value between 0 to 1 for the permitted transitions and 0 for the forbidden ones (Figure 7.4*).* The matrix has a dual purpose: it specifies allowed transitions and provides initial values of *transition intensities.* Equations 7.1 – 7.4 provides the mathematical formulations and detailed description of processes that are used to fit a model to the data.

$$Q = \begin{array}{c} \\ [BL\ (1)] \\ [Neither\ (2)] \\ [Hosp\ (3)] \\ [Dead\ (4)] \end{array} \begin{array}{c} \begin{array}{cccc} [BL\ (1)] & [Neither]\ (2) & [Hosp](3) & [Dead\ (4)] \end{array} \\ \begin{bmatrix} -(q_{1,2} + q_{1,3} + q_{1,4}) & q_{1,2} & q_{1,3} & q_{1,4} \\ 0 & -(q_{2,3} + q_{2,4}) & q_{2,3} & q_{2,4} \\ 0 & q_{3,2} & -(q_{3,2} + q_{3,4}) & q_{3,4} \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Figure 7.4: Q matrixTo allow a multistate model to be fitted to data, figure 7.3 was translated into a square matrix (known as the *Q matrix)*, taking a value between 0 to 1 for the permitted transitions and 0 for the forbidden ones. The matrix is denoted by *Q* and intensities for the corresponding transitions are denoted by $q_{j,k}$. Here, BL, Neither, Hosp and Dead states are represented by 1,2,3 and 4 respectively.

As an example, the entry $q_{1,2}$ represents the intensity for the transition from BL to Neither. Remaining in the same sate in the next cycle is represented by a minus sign (e.g., $1-(q_{1,2} + q_{1,3} + q_{1,4})$). The transition intensity for remaining in the same state is obtained by subtracting the total of all other probabilities in the same row from 1. For abbreviations see figure 7.3.

7.2.6 Study outcomes

Multi-state models (MSM) analyse situations where individuals in a population can transition between a number of states depending upon continuing risk (the intensities for transition being assumed to be constant with time)[7] (Meira-Machado et al., 2009; Kazmi et al., 2022). MSMs provide a general and flexible framework that extends beyond the Kaplan-Meier estimator and Cox models (Le-Rademacher et al., 2022). Unlike these models which typically assume a one-directional progression towards an event, my model allow subjects to move continuously between states including both "backwards" (from *[Hosp]* to *[Neither])* and "forwards" transitions (from *[Neither]* to *[Hosp]* or *[Death]*).

I used features of multistate modelling to derive summary information from the model (Figure 7.2). I used the model to estimate the *transition intensities*[8] (equivalent to relative risk in Cox modelling) between states with 95% confidence intervals (CI). The CI represent the precision (error) of the predicted intensities.

The multistate models can be used to simulate the progression of individuals through various health states until they reach the endpoint of death, which is often the final absorbing

[7] This refers to a key assumption often used in the modelling of processes with MSMs. Assuming these intensities are constant over time means that these rates are not expected to change throughout the period of observation.

[8] The *transition intensities* are the instantaneous *hazards of transition*, representing the instantaneous risk of moving from a current state (*j)* to a subsequent state (*k).* The summary information was calculated with and without adjusting for covariates.

state (often *[Death]*) in this model. I derived the time spent in a given state in total before death as predicted by the model (referred as *total length of stay* in MSM modelling*)*.

7.2.7 Model validation

The analysis was conducted in stages: first, an unadjusted model was constructed using all patients with complete data at baseline (cohort a). The model was developed from the observed transitions in the first two cycles (Figure 7.5); and the predictions of the model for the following three years (3^{rd} to 5^{th} cycles) was compared with the observed data. This unadjusted model was independent of any possible covariates.

I then validated the model's performance using a second cohort of patients, cohort b. The second cohort consisted of patients in whom one or more of the potential clinical covariates were missing (see below) and hence could not be included in the multivariable model. Similar to the approach used for cohort (a), the model's predictions were compared with observed data (transitions) of cohort (b) across the first five yearly transitions.

Finally, I introduced clinical covariates to develop a second model. Fitting an MSM with covariates can be complex and computationally intensive due to a potentially large number of parameters[9] (Jackson, 2019). To get a parsimonious model and to avoid the risk of over-

[9] In the study, the term 'parameters' refers to the coefficients, constants, equations or functions used in Markov and multi-state modelling to represent relationships between variables. These parameters are estimated from the data and can provide insight into the nature and strength of relationships between variables.

fitting, I followed a step-by-step procedure. I started with univariable analysis using all the variables in Table 8.1. The models were reduced by backward selection, eliminating those with no impact on transition intensities: the final model contained only: age, sex, NT-proBNP, NYHA class, haemoglobin, urea, albumin, a loop diuretic (yes/no). Because most of the covariates in the model (other than age and sex) are related to severity of cardiovascular disease, I also compared the performance of the two models against the end-point of cardiovascular mortality.

7.2.8 Model diagnostics and assessment of goodness-to-fit

The *likelihood* is the probability that a given set of parameters defining a model are correct, given the observed data. I used the *likelihood ratio* test (defined as $-2\ln(L_s(\hat{\theta}) - L_g(\hat{\theta}))$, where $L_s(\hat{\theta})$ is the likelihood of the unadjusted model and $L_g(\hat{\theta})$ is the likelihood of the covariate-adjusted model) to show which of the two models better fitted the observed data. The lower the likelihood ratio, the better the fit of the multivariable model. Significance was set at an arbitrary level of 5% (two − tailed). The aim was to get a parsimonious model that accurately describes real events.

A formal test of goodness-of-fit was performed by plotting observed and expected percentages against time (t) (Meira-Machado et al., 2009; Jackson, 2019). A Z-score was used to standardise variables to a distribution with a mean of zero and a standard deviation of 1.

Confidence intervals were estimated on 1000 bootstraps (bootstrapping is sampling *with* replacement).

Preliminary analyses used Stata software. The multistate model was determined using R (4.2.2) running the *MSM* and *tdc.msm* packages (Meira-Machado et al., 2007; Jackson, 2019). NT-proBNP value was log-10 transformed.

7.3 Discussion

In this chapter, I developed a multistate risk model to describe the progression of chronic heart failure (CHF). It included detailed mathematical formulations and a comprehensive description of the processes for fitting the model to the data. Various clinical covariates are incorporated to examine their influence on transitions between mutually exclusive clinical states. The model provided the valuable insights into disease progression which are detailed in next chapter.

Chapter 8 Insights from multistate model of chronic HF

The results derived from the multistate risk model (developed in Chapter 7) are presented. I analyse the key findings and discuss their significance.

8.1 Descriptive characteristics

Table 8.1 shows the demographic and clinical characteristics of the patients with (a) complete data (n = 4989) and (b) incomplete data (n =1830). The patients in (a) were slightly older and were more likely to be men compared to cohort (b). Cohort (a) was classified into different diagnostic categories as outlined in section 8.1.1. Their baseline demographic and clinical characteristics are also detailed in the table. The distribution and proportion of patients following the first two transitions (between baseline to 12 months and 12 months to 24 months) is shown in Table 8.2. Mortality is low in the HeFPEF 125-399 diagnostic category, with only 4% and 10 % of patients in this group having died by the end of the first and second cycles respectively. Transitions from the first two cycles (used to develop the model) in Markovian structure are shown in Figure 8.1.

There were 20313 transitions during the five year follow up in cohort (a), of which 10% (n = 2051; 41% of the patients) were deaths, and 27% (n = 5544) were hospitalisations. (In all the other "transitions", a patient neither died nor was hospitalised).

Variable	Cohort a (complete data)					Cohort b (incomplete data)		*P*
	Missing (n (%))	Total	HeFREF	HeFPEF		Missing (n (%))	Total	
				>=400	125-399			
Total number of patients N (%)		4989 (100)	2180 (44)	1909 (38)	900 (18)		1830 (100)	
Demographics								
Age (years)		74.6 [67.1, 80.6]	72.5 [64.2, 78.6]	77.6 [71.5, 82.8]	72.6 [65.7, 79.0]		73.6 [65.7, 80.0]	<0.001
<75		2464 (49.3)	1274 (58.4)	678 (35.5)	512 (56.9)		989 (54.0)	0.001
≥75		2525 (50.6)	906 (41.6)	1231 (64.5)	388 (43.1)		841 (46.0)	0.001
Men (%)		3060 (61.3)	1619 (74.3)	1008 (52.8)	433 (48.1)		1000 (54.6)	<0.001
BMI (kg/m2)		28.0 [24.6, 31.9]	27.3 [24.0, 30.9]	28.1 [24.7, 32.2]	29.6 [26.1, 33.5]	67 (4)	27.5 [24.3, 31.8]	0.07
NYHA class (%)								
I-II		3443 (69.0)	1365 (62.6)	1316 (68.9)	762 (84.7)		1365 (74.6)	<0.001
III-IV		1546 (31.0)	815 (37.4)	593 (31.1)	138 (15.3)		465 (25.4)	<0.001
Systolic BP (mmHg)		138.0 [120.0, 157.0]	128.0 [113.0, 145.0]	142.0 [126.0, 162.0]	149.0 [133.0, 165.2]		141.0 [124.0, 161.0]	<0.001
Diastolic BP (mmHg)		78.0 [69.0, 88.0]	76.0 [67.0, 86.0]	78.0 [68.8, 89.0]	80.0 [72.0, 90.0]		80.0 [70.0, 89.0]	<0.001
Left ventricular systolic dysfunction								
LV Impairment								
None		1692 (33.9)	0 (0.0)	1058 (55.4)	634 (70.4)		982 (53.6)	<0.001
Trivial		390 (7.8)	0 (0.0)	291 (15.2)	99 (11.0)		144 (7.9)	0.984
Mild		727 (14.6)	0 (0.0)	560 (29.3)	167 (18.6)		199 (10.9)	<0.001
Worse		2180 (43.7)	2180 (100.0)	0 (0.0)	0 (0.0)		505 (27.6)	<0.001
Findings on electrocardiogram								
Heart rate (bpm)		72.0 [62.0, 84.0]	74.0 [64.0, 87.0]	72.0 [62.0, 85.0]	67.0 [59.0, 78.0]	227 (12)	72.0 [63.0, 84.0]	0.899

Heart rhythm (Sinus %)	3020 (60.5)	1347 (61.8)	861 (45.1)	812 (90.2)		1142 (62.4)	0.169
QRS (ms)	100.0 [88.0, 124.0]	114.0 [98.0, 144.0]	96.0 [84.8, 112.0]	92.0 [82.0, 100.0]	286 (16)	98.0 [86.0, 112.5]	<0.001
Blood test							
Haemoglobin (g/dL)	13.3 [12.1, 14.4]	13.5 [12.3, 14.6]	12.8 [11.7, 14.1]	13.5 [12.6, 14.4]	1052 (58)	13.2 [12.0, 14.4]	0.49
Sodium (mmol/L)	139.0 [136.0, 140.0]	138.0 [136.0, 140.0]	139.0 [137.0, 140.0]	139.0 [137.0, 141.0]	865 (47)	139.0 [136.0, 140.0]	0.417
Potassium (mmol/L)	4.3 [4.1, 4.7]	4.4 [4.1, 4.7]	4.3 [4.0, 4.7]	4.3 [4.1, 4.6]	879 (48)	4.3 [4.0, 4.7]	0.237
Urea (mmol/L)	6.8 [5.2, 9.3]	7.1 [5.4, 9.8]	7.1 [5.4, 9.8]	5.8 [4.7, 7.3]	867 (47)	6.8 [5.2, 9.7]	0.58
Creatinine (umol/L	99.0 [82.0, 125.0]	105.0 [86.0, 132.0]	100.0 [82.0, 129.0]	87.0 [74.0, 102.0]	871 (48)	100.0 [82.0, 126.0]	0.508
Albumin (g/L)	38.0 [35.0, 40.0]	38.0 [35.0, 40.0]	37.0 [35.0, 39.0]	38.0 [36.0, 40.0]	961 (53)	37.0 [34.0, 40.0]	0.001
NT-proBNP (ng/L)	1099.4 [422.8, 2604.8]	1750.6 [748.1, 3914.8]	1307.0 [750.0, 2479.6]	226.8 [167.1, 296.1]	1232 (67)	752.9 [310.2, 2025.0]	<0.001
Heart failure medication							
Beta-blocker (%)	2590 (51.9)	1529 (70.1)	833 (43.6)	228 (25.3)		520 (28.4)	
Low dose	1477 (57)	931 (60.9)	439 (52.7)	107 (46.9)		307 (59)	<0.001
High dose	1113 (43)	598 (39.1)	394 (47.3)	121 (53.1)		213 (41)	<0.001
ACE/ARB (%)	3819 (76.5)	1972 (90.5)	1317 (69.0)	530 (58.9)		986 (53.9)	<0.001
MRA (%)	1299 (26.0)	893 (41.0)	336 (17.6)	70 (7.8)		268 (14.6)	<0.001

Table 8.1: Baseline demographic and clinical characteristics for two different cohortsa) patients with completed data (N = 4989) and their diagnostic categories (CHF phenotypes: HeFREF and HeFPEF (>=400 and 125-399)). b) patients with incomplete data (N = 1830). *P* values are for difference between total patients in cohorts, model was developed on cohort (a) and validated on cohort (b).
*Continuous variables are presented as median (interquartile range), whereas categorical variables are expressed as numbers and percentage.
P value indicates significance at the 0.05 level. *NT-proBNP only became a clinical service during the course of the data collection. See section 4 of chapter 4 for abbreviations.

From BL to end of 1st cycle (N = 4989)				
Variable	**Neither (N (%))**	**Hosp (N (%))**	**Dead (N (%))**	**Total N (%)**
	2714 (54)	1699 (34)	576 (12)	4989 (100)
Age (years)				
< 75	1375 (56)	908 (37)	181 (07)	2464 (100)
≥ 75	1339 (53)	791 (31)	395 (16)	2525 (100)
Sex				
Female	1102 (57)	613 (32)	214 (11)	1929 (100)
Male	1612 (53)	1086 (35)	362 (12)	3060 (100)
Diagnostic categories				
HeFREF	1117 (51)	782 (36)	281 (13)	2180 (100)
HeFPEF ≥ 400	987 (52)	661 (35)	261 (14)	1909 (100)
HeFPEF 125-399	610 (68)	256 (28)	34 (4)	900 (100)
From end of 1st cycle to end of 2nd cycle (N = 4413)				
Variable	**Neither (N (%))**	**Hosp (N (%))**	**Dead (N (%))**	**Total N (%)**
	2898 (66)	1168 (26)	406 (9)	4413 (100)
Age (years)				
< 75	1488 (65)	532 (23)	263 (12)	2283 (100)
≥ 75	1144 (54)	466 (22)	520 (24)	2130 (100)
Sex				
Female	1068 (62)	371 (22)	276 (16)	1715 (100)
Male	1564 (58)	627 (23)	507 (19)	2698 (100)
Diagnostic categories				
HeFREF	1069 (56)	464 (24)	366 (19)	1899 (100)
HeFPEF ≥ 400	939 (57)	378 (23)	331 (20)	1648 (100)
HeFPEF 125-399	624 (72)	156 (18)	86 (10)	866 (100)

Table 8.2: Total number of observed transitions among states.

Stratified by selective variables *Data is expressed as numbers and percentage (round to 2 digits). Abbreviation: BL; baseline, Hosp; Hospitalisation; HF, Heart failure; HeFREF, HF with reduced ejection fraction; HeFPEF, HF with preserved ejection fraction; N, total number; %, percentage.

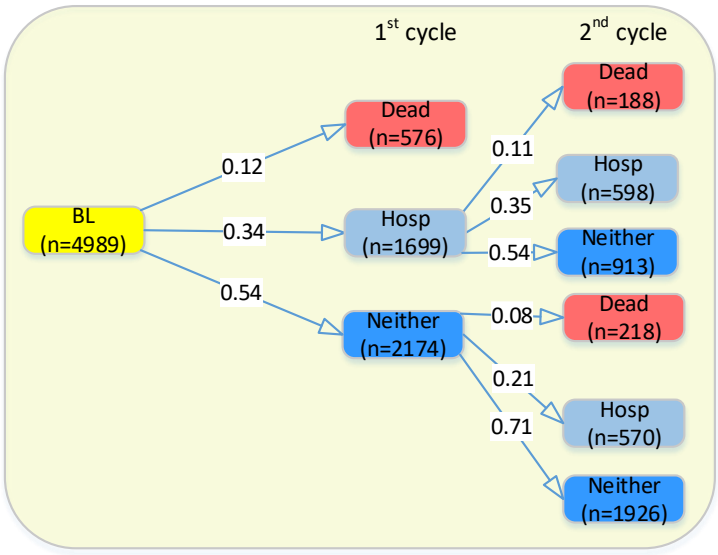


Figure 8.1: Transitions from the first two cycles (cohort (a))These cycles were used to develop the model.

8.1.1 Unadjusted model

Table 8.3 shows the predictions made by the unadjusted model (M1) and the deviations from the observed data. As examples, the predicted transition intensity from hospital to death is 0.11, meaning that of 100 patients who are hospitalised at the end of one cycle, 11 will have died by the end of the next. Similarly, the transition intensity from [*Neither*] to death of 0.08 means that of 100 patients who have not been hospitalised at the end of one cycle, 8 will have died by the end of the next. The differences between model predictions and observed events are also shown. The majority of observed transition probability fall within the CI of model (M1) predictions. Meaning errors are zero. This close alignment between predicted and

observed values underscores the model's effectiveness in capturing the dynamics of disease progression and the risks associated with hospitalisation.

For a hospitalised patient, recovery (defined as the ratio of the intensities from [*Hosp*] to [*Neither*] and [*Hosp*] to [*Dead*]) was approximately 5 times more likely than death (0.54/0.11). For further detail and interpretation of the data, refer to the table legend.

Models prediction in the derivation cohort (for those with complete data: N = 4989)

State-to-state transition	Unadjusted model (M1)	Observed transition intensities for each cycle					Error (predicted (M1) - observed)					Multivariable model (M2)
	TI (95% CI)	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	5th	TI (95% CI)
BL - Neither	0.54 (0.52, 0.57)	no further transitions from BL after first cycle					NA					0.59 (0.54, 0.64)
BL - Hosp	0.34 (0.32, 0.36)											0.34 (0.32, 0.35)
BL - Dead	0.12 (0.10, 0.13)											0.08 (0.07, 0.09)
Neither - Neither	0.71 (0.69, 0.73)			0.73	0.71	0.72			-0.02	0.00	-0.01	0.72 (0.70, 0.74)
Neither - Hosp	0.21 (0.19, 0.23)			0.2	0.21	0.20			0.01	0.00	0.01	0.21 (0.19, 0.23)
Neither - Dead	0.08 (0.07, 0.09)			0.08	0.08	0.08			0.00	0.00	0.00	0.07 (0.06, 0.08)
Hosp - Neither	0.54 (0.50, 0.57)			0.49	0.52	0.49			0.05	0.02	0.04	0.54 (0.51, 0.58)
Hosp - Hosp	0.35 (0.31, 0.39)			0.38	0.34	0.35			-0.03	0.01	0.01	0.38 (0.34, 0.42)
Hosp - Dead	0.11 (0.10, 0.13)			0.13	**0.14**	0.16			-0.02	-0.03	-0.05	0.08 (0.07, 0.10)

Model performance in the validation cohort (for those in whom one or more of the covariates were missing: N = 1830)

State-to-state transition	Unadjusted model (M1)	Observed transition intensities for each cycle					Error (predicted (M1) - observed)				
	TI (95% CI)	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	5th
BL - Neither	0.54 (0.52, 0.57)	0.57	no further transitions from BL after first cycle				-0.03	NA			
BL - Hosp	0.34 (0.32, 0.36)	0.26					0.08				
BL - Dead	0.12 (0.10, 0.13)	0.17					-0.05				
Neither - Neither	0.71 (0.69, 0.73)	NA	0.78	0.80	0.77	0.79	NA	-0.07	-0.09	-0.06	-0.08
Neither - Hosp	0.21 (0.19, 0.23)		0.15	0.13	0.17	0.16		0.06	0.08	0.04	0.05

Neither - Dead	0.08	(0.07, 0.09)		0.07	0.07	0.05	0.06		0.01	0.01	0.03	0.02
Hosp - Neither	0.54	(0.50, 0.57)		0.53	0.56	0.55	0.50		0.01	-0.02	-0.01	0.04
Hosp - Hosp	0.35	(0.31, 0.39)		0.32	0.31	0.31	0.37		0.03	0.04	0.04	-0.01
Hosp - Dead	0.11	(0.10, 0.13)		0.15	0.14	0.14	0.13		-0.04	-0.03	-0.03	-0.02

Table 8.3: Predicted and observed probabilities

Table shows he predicted transition probabilities of unadjusted model (M1) and multivariable model (M2) and model validation using the observed probabilities from the raw data of following cohort (a) and (b) In the first half of the table the column 2 (light grey) shows the predictions made by the unadjusted model (along with 95% CI). The observed transition probabilities are shown in the central columns (darker grey). The model was derived using observed data for the first two cycles. All patients started in [BL] state and transition from this state was unidirectional to one of three other states. Data is expressed as transition probabilities with CI (rounded to 2 digits). The columns shaded green to red (by size of error) show the precision of the unadjusted model. Minus (-) signs mean underestimation of model; the others are overestimates. Note: The majority of observed transition probability fall within the CI of model (M1) predictions. Meaning errors are zero. For example, the model predicts the transition intensity of from [*Hosp*] to death is 0.11 with 95% CIs (0.10 - 0.13)). The observed probability of this transition in the fourth cycle of cohort (a) is 0.14 (bold) which falls only 0.01 outside of CI of the model prediction.
The right-hand column (yellow) shows the predictions made by the multivariable model after adjusting covariates. Transition probability is equivalent to the *transition intensity* (TI) of a specific transition during a single cycle in an MSM.

The second half of the table shows the performance of unadjusted model (M1) (as mentioned above) against the probabilities obtained (central columns (darker grey)) from the validation cohort. To show the error for 1st and 2nd annual cycles, the transitions from BL to 1st year and 1st to 2nd year were also included. The columns shaded green to red (by size of error) show the precision of the unadjusted model.

Abbreviation: BL; baseline, Hosp; Hospitalisation, TI; Transition intensities, CI; Confidence Interval. NA; not applicable. The likelihood ratio (-2*log-likelihood) of the unadjusted model was: 28048.21 and for the adjusted model was: 27140.98

8.1.2 Univariable analysis

Univariable analysis was performed using all the variables in Table 8.1. The models were reduced by backward selection, eliminating those with no impact on transition intensities: the final model contained only: age, sex, NT-proBNP, NYHA class, haemoglobin, urea, albumin, a loop diuretic (yes/no). For each covariate, a separate independent variable is fitted as shown in Table 8.4. Example interpretation, age ≥ 75 years [*BL*] to [*Dead*] there is 13% chance of individual died and relative risk of dying is doubled compare to those < 75 years old. The rest follow suit.

Covariate	State transition	TI (CI = 95%)		HR (95% CI)	
Age ≥75 (years)	**BL - Neither**	0.53	(0.49, 0.58)	0.93	(0.87, 0.99)
	BL - Hosp	0.34	(0.33, 0.37)	0.97	(0.89, 1.05)
	BL - Dead	0.13	(0.12, 0.14)	2.21	(1.93, 2.54)
	Neither - Neither	0.73	(0.71, 0.75)	0.88	(0.82, 0.95)
	Neither - Hosp	0.20	(0.18, 0.21)	1.21	(1.05, 1.40)
	Neither - Dead	0.07	(0.06, 0.08)	2.52	1.97, 3.22)
	Hosp - Neither	0.56	(0.53, 0.59)	0.91	(0.81, 1.02)
	Hosp - Hosp	0.33	(0.33, 0.36)	0.94	(0.81, 1.09)
	Hosp - Dead	0.11	(0.10, 0.13)	2.57	(1.98, 3.34)
Sex (male)					
	BL - Neither	0.52	(0.48, 0.57)	0.93	(0.87, 0.99)
	BL - Hosp	0.34	(0.33, 0.35)	1.13	(1.03, 1.23)
	BL - Dead	0.14	(0.13, 0.15)	1.08	(0.95, 1.24)
	Neither - Neither	0.72	(0.71, 0.74)	0.95	(0.88, 1.03)
	Neither - Hosp	0.20	(0.18, 0.21)	1.12	(0.95, 1.30)
	Neither - Dead	0.08	(0.07, 0.09)	1.35	(1.06, 1.71)
	Hosp - Neither	0.56	(0.53, 0.59)	0.94	(0.83, 1.05)
	Hosp - Hosp	0.32	(0.28, 0.35)	1.14	(0.98, 1.32)
	Hosp - Dead	0.12	(0.11, 0.14)	1.00	(0.78, 1.28)
NT-proBNP (log10)					
	BL - Neither	0.56	(0.51, 0.61)	0.76	(0.71, 0.81)
	BL - Hosp	0.35	(0.34, 0.37)	1.22	(1.12, 1.33)
	BL - Dead	0.09	(0.08, 0.10)	**4.28**	**(3.71, 4.93)**

	Neither - Neither	0.71	(0.69, 0.73)	0.86	(0.79, 0.93)
	Neither - Hosp	0.21	(0.20, 0.23)	1.30	(1.12, 1.51)
	Neither - Dead	0.08	(0.07, 0.09)	2.40	(1.90, 3.05)
	Hosp - Neither	0.56	(0.53, 0.60)	0.82	(0.73, 0.93)
	Hosp - Hosp	0.34	(0.30, 0.38)	1.09	(0.94, 1.27)
	Hosp - Dead	0.10	(0.08, 0.12)	2.66	(2.05, 3.45)
NYHA (III/IV)					
	BL - Neither	0.54	(0.50, 0.59)	0.71	(0.66, 0.77)
	BL - Hosp	0.34	(0.32, 0.35)	1.32	(1.21, 1.45)
	BL - Dead	0.12	(0.11, 0.13)	2.92	(2.56, 3.34)
	Neither - Neither	0.72	(0.70, 0.74)	0.89	(0.81, 0.98)
	Neither - Hosp	0.20	(0.18, 0.21)	1.16	(0.98, 1.38)
	Neither - Dead	0.08	(0.07, 0.09)	2.35	(1.86, 2.97)
	Hosp - Neither	0.56	(0.53, 0.60)	0.79	(0.69, 0.86)
	Hosp - Hosp	0.32	(0.28, 0.35)	1.22	(1.05, 1.42)
	Hosp - Dead	0.12	(0.10, 0.13)	1.74	(1.36, 2.22)
Haemoglobin					
	BL - Neither	0.52	(0.48, 0.57)	1.05	(1.03, 1.07)
	BL - Hosp	0.36	(0.34, 0.37)	0.98	(0.95, 1.00)
	BL - Dead	0.12	(0.11, 0.13)	0.75	(0.72, 0.78)
	Neither - Neither	0.71	(0.69, 0.73)	1.03	(1.01, 1.06)
	Neither - Hosp	0.21	(0.19, 0.23)	0.94	(0.89, 0.98)
	Neither - Dead	0.08	(0.07, 0.09)	0.87	(0.81, 0.94)
	Hosp - Neither	0.56	(0.53, 0.59)	1.05	(1.02, 1.09)
	Hosp - Hosp	0.32	(0.29, 0.36)	0.95	(0.91, 1.00)
	Hosp - Dead	0.12	(0.10, 0.13)	0.86	(0.80, 0.92)
Urea					
	BL - Neither	0.52	(0.47, 0.56)	0.98	(0.97, 0.99)
	BL - Hosp	0.36	(0.34, 0.37)	1.01	(1.00, 1.01)
	BL - Dead	0.13	(0.12, 0.14)	1.08	(2.56, 3.34)
	Neither - Neither	0.71	(0.69, 0.73)	0.98	(0.97, 0.99)
	Neither - Hosp	0.21	(0.20, 0.23)	1.02	(1.01, 1.04)
	Neither - Dead	0.08	(0.07, 0.09)	1.07	(1.06, 1.09)
	Hosp - Neither	0.56	(0.53, 0.59)	0.99	(0.97, 1.00)
	Hosp - Hosp	0.32	(0.28, 0.36)	1.00	(0.99, 1.02)
	Hosp - Dead	0.12	(0.10, 0.14)	1.05	(1.03, 1.07)
Albumin					
	BL - Neither	0.53	(0.48, 0.58)	1.03	(0.97, 0.99)
	BL - Hosp	0.36	(0.34, 0.37)	0.99	(1.00, 1.01)
	BL - Dead	0.12	(0.11, 0.13)	0.85	(2.56, 3.34)
	Neither - Neither	0.71	(0.69, 0.73)	1.01	(0.97, 0.99)
	Neither - Hosp	0.21	(0.19, 0.23)	1.00	(0.97, 1.02)
	Neither - Dead	0.08	(0.07, 0.09)	0.91	(0.88, 0.94)
	Hosp - Neither	0.56	(0.53, 0.59)	1.02	(1.01, 1.04)
	Hosp - Hosp	0.33	(0.29, 0.36)	0.99	(0.97, 1.01)
	Hosp - Dead	0.11	(0.10, 0.13)	0.91	(0.88, 0.94)
Loop diuretic (yes)					
	BL - Neither	0.54	(0.49, 0.58)	0.80	(0.75, 0.85)

BL - Hosp	0.34	(0.32, 0.35)	1.28	(1.17, 1.40)
BL - Dead	0.13	(0.12, 0.14)	2.39	(2.03, 2.82)
Neither - Neither	0.73	(0.71, 0.74)	0.85	(0.79, 0.91)
Neither - Hosp	0.20	(0.18, 0.21)	1.46	(1.25, 1.70)
Neither - Dead	0.08	(0.07, 0.09)	2.83	(2.14, 3.74)
Hosp - Neither	0.56	(0.53, 0.60)	0.88	(0.88, 0.99)
Hosp - Hosp	0.32	(0.29, 0.36)	1.06	(0.91, 1.23)
Hosp - Dead	0.11	(0.10, 0.13)	2.13	(1.56, 2.92)

Table 8.4: Cox Markov model for independent variablesTransition trajectories and unadjusted hazard ratios (HRs) for independent variables Abbreviations: TI, transition intensities; CI, confidence interval; HR, hazard ratio, NYHA, New York Heart Association; BL, baseline; Hosp, Hospitalisation; Dead, death. Example interpretation, BL to dead: for every increase in log10 NTproBNP HR increase 4 folds (**bold**).

8.1.3 Multivariable model

The right-hand column (yellow) of Table 8.3 shows the predictions made by the multivariable model (M2) after adjusting covariates. The multivariable model under-estimated the intensity of transitions from any state to death. As an example, for the transition from [*Hosp*] to [*Dead*] the difference between predicted and observed events at cycle 3, 4 and 5 were -0.02, 0.03 and 0.05 for the unadjusted model, and -0.05, -0.06 and -0.08 for the multivariable model, respectively. A likelihood ratio test is performed as defined in the model diagnostics section. The model gave a slightly better fit compared to the unadjusted model ($p < 0.001$), and the likelihood ratio test was lower (27,140 vs 28,048). Table 8.5 shows the performance of the model against the end-point of cardiovascular mortality for cohort (a), the multivariable model perform little better (smaller LLR) than the unadjusted model in predicting the endpoint death when patients transit from [*Hosp*] to [*Dead*] state.

State-to-state transition	Adjusted model (M2)	Observed transition intensities for each cycle					Error (predicted - observed)				
	TI (95% CI)	1st CV	2nd CV	3rd CV	4th CV	5th CV	1st	2nd	3rd	4th	5th
BL - Neither	0.59 (0.54, 0.64)										
BL - Hosp	0.34 (0.32, 0.35)		no further transitions from BL after first cycle					NA			
BL - Dead	0.08 (0.07, 0.09)	0.08					0.00				
Neither - Neither	0.72 (0.70, 0.74)										
Neither - Hosp	0.21 (0.19, 0.23)										
Neither - Dead	0.07 (0.06, 0.08)	NA	0.05	0.04	0.04	0.04	NA	-0.02	-0.03	-0.03	-0.03
Hosp - Neither	0.54 (0.51, 0.58)										
Hosp - Hosp	0.38 (0.34, 0.42)										
Hosp - Dead	0.08 (0.07, 0.10)		0.06	0.07	0.08	0.08		-0.02	-0.01	0.00	0.00

Table 8.5: Multivariable models against the end-point of CV mortality of cohort (a)

The risk of transitioning can be seen by the hazard ratios in multivariable analysis (Table 8.6)). In older patients (≥ 75 years), the risk of moving from [*Hosp*] state to death more than doubled (2.20) with increasing age, adjusted for the other covariates. An observation is the risk of going from [*BL*] to [*Hosp*] for older people was less than 1 (0.85); this reflects the selection bias where older people move from [*BL*] to death in the first cycle.

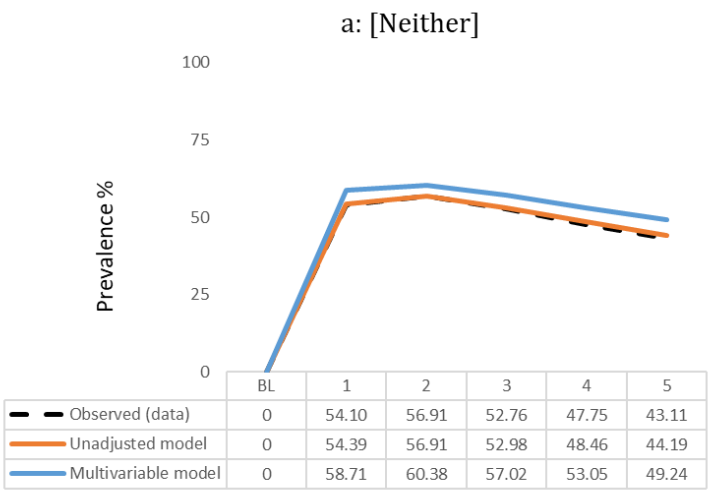
To better understand the influence of these variables on the state to state transitions. I standardised the variables to Z-scores (mean = 0, SD = 1), all on the same scale. This allows us to compare directly which of those independent variables have the biggest effect on transition. Looking at the Z-score hazard ratios moving from [*BL*] to death, NT-proBNP has the higher hazard (1.70) then age, NYHA and urea, with hazards of 1.19, 1.18 and 1.10, respectively. The strongest influence in the transition to death from [*Neither*] or from [*Hosp*] is age (1.31 and 1.48, respectively). Sex and NYHA have higher HRs (1.11 and 1.09, respectively) for Re-hospitalisation. Moving to [*Hosp*] state from [*BL*], the ordering is different, with high NYHA (III/IV) (1.09), NTproBNP (1.07) and male sex (1.06). [*Neither*] moving to [*Hosp*], NT-proBNP, Albumin, loop diuretic have the strongest influence (with same Z-score (1.10)) and male gender (1.09). There is little influence of urea and age (1.02).

Cox Markov model								
Transitions	**Age-group**	**Sex**	**NT-proBNP**	**NYHA**	**Haemoglobin**	**Urea**	**Albumin**	**Loop diuretic**
	≥ 75 years	**Male**	**(ng/L)**	**III-IV**	**(g/dL)**	**(mmol/L)**	**(g/L)**	**Yes**
BL - Neither	1.06 (0.98, 1.15)	0.91 (0.84, 0.99)	0.79 (0.72, 0.85)	0.79 (0.72, 0.86)	1.04 (1.01, 1.06)	0.99 (0.98, 1.00)	1.01 (1.00, 1.02)	1.00 (0.92, 1.09)
BL - Hosp	**0.85 (0.77, 0.94)**	1.12 (1.01, 1.24)	1.11 (1.00, 1.22)	1.21 (1.09, 1.35)	0.98 (0.95, 1.01)	0.99 (0.98, 1.01)	1.01 (1.00, 1.03)	1.06 (0.95, 1.19)
BL - Dead	1.42 (1.19, 1.70)	1.10 (0.93, 1.31)	2.19 (1.84, 2.91)	1.46 (1.23, 1.74)	0.91 (0.87, 0.96)	1.02 (1.01, 1.04)	0.94 (0.92, 0.96)	1.10 (0.88, 1.38)
Neither - Neither	0.94 (0.86, 1.04)	0.91 (0.83, 1.01)	0.93 (0.84, 1.03)	0.94 (0.84, 1.05)	1.03 (0.99, 1.06)	0.99 (0.98, 1.00)	1.00 (0.99, 1.01)	0.92 (0.83, 1.02)
Neither - Hosp	1.03 (0.87, 1.23)	1.17 (0.98, 1.40)	1.15 (0.96, 1.37)	1.02 (0.87, 1.24)	0.93 (0.87, 0.98)	1.01 (0.99, 1.03)	1.03 (1.00, 1.05)	1.23 (1.01, 1.50)
Neither - Dead	1.71 (1.27, 2.30)	1.43 (1.06, 1.92)	1.41 (1.07, 1.87)	1.42 (1.07, 1.89)	1.00 (0.92, 1.09)	1.03 (1.01, 1.06)	0.94 (0.91, 0.98)	1.63 (1.13, 2.35)
Hosp - Neither	0.90 (0.78, 1.03)	0.86 (0.74, 0.99)	0.90 (0.78, 1.04)	0.81 (0.70, 0.93)	1.04 (0.99, 1.08)	1.00 (0.98, 1.01)	1.01 (0.99, 1.03)	1.03 (0.88, 1.19)
Hosp - Hosp	0.94 (0.79, 1.11)	1.21 (1.01, 1.45)	1.00 (0.84, 1.19)	1.19 (1.00, 1.41)	0.95 (0.90, 1.00)	1.00 (0.98, 1.01)	1.00 (0.98, 1.02)	0.94 (0.78, 1.13)
Hosp - Dead	2.20 (1.60, 3.03)	1.13 (0.83, 1.55)	1.72 (1.25, 2.34)	1.51 (1.12, 2.03)	1.01 (0.92, 1.11)	1.03 (1.00, 1.05)	0.95 (0.92, 0.99)	1.24 (0.84, 1.83)
(b) Z-score hazard of transitioning								
BL - Neither	1.03 (0.99, 1.07)	0.96 (0.92, 0.99)	0.85 (0.80, 0.90)	0.91 (0.87, 0.94)	**1.06 (1.01, 1.10)**	0.98 (0.93, 1.02)	1.05 (1.00, 1.09)	1.00 (0.96, 1.04)
BL - Hosp	**0.92 (0.88, 0.97)**	1.06 (1.00, 1.11)	**1.07 (1.00, 1.15)**	**1.09 (1.04, 1.14)**	0.96 (0.91, 1.00)	0.97 (0.92, 1.03)	1.04 (0.96, 1.10)	1.03 (0.97, 1.09)
BL - Dead	1.19 (1.09, 1.30)	1.05 (0.96, 1.14)	**1.70 (1.52, 1.91)**	1.18 (1.09, 1.27)	0.86 (0.79, 0.93)	1.10 (1.04, 1.16)	0.80 (0.69, 0.81)	1.05 (0.94, 1.17)
Neither - Neither	0.97 (0.93, 1.02)	0.96 (0.91, 1.01)	0.95 (0.89, 1.01)	0.98 (0.93, 1.03)	**1.04(0.99, 1.10)**	0.95 (0.99, 1.01)	1.00 (0.95, 1.05)	0.96 (0.92, 1.01)
Neither - Hosp	1.02 (0.93, 1.11)	1.09 (1.00, 1.17)	**1.10 (0.97, 1.24)**	1.01 (0.93, 1.05)	0.88 (0.79, 0.97)	1.03 (0.93, 1.13)	**1.10 (0.99, 1.21)**	**1.10 (1.01, 1.22)**
Neither - Dead	**1.31 (1.31, 1.52)**	1.19 (1.03, 1.38)	1.26 (1.04, 1.53)	1.17 (1.04, 1.33)	1.00 (0.86, 1.16)	1.15 (1.03, 1.29)	0.79 (0.68, 0.91)	1.28 (1.07, 1.54)
Hosp - Neither	0.95 (0.89, 1.05)	0.93 (0.87, 1.00)	0.93 (0.86, 1.04)	0.91 (0.86, 0.97)	**1.06 (0.98, 1.14)**	0.99 (0.92, 1.07)	1.04 (0.97, 1.12)	1.01 (0.94, 1.09)
Hosp - Hosp	0.97 (0.89, 1.06)	**1.10 (1.01, 1.20)**	1.00 (0.89, 1.12)	**1.09 (1.01, 1.17)**	0.90 (0.83, 0.99)	0.97 (0.89, 1.06)	1.00 (0.91, 1.09)	0.97 (0.88, 1.06)
Hosp - Dead	**1.48 (1.26, 1.74)**	1.07 (0.92, 1.24)	1.41 (1.14, 1.74)	1.20 (1.05, 1.37)	1.02 (0.87, 1.19)	1.11 (1.01, 1.24)	0.82 (0.71, 0.95)	1.12 (0.92, 1.35)

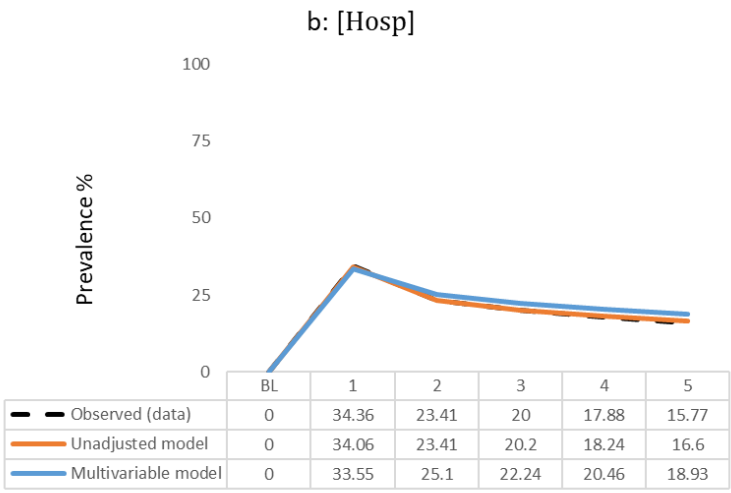
Table 8.6: Multivariable Cox Markov model with Z-Score Table shows multivariable hazard ratio (HR) and Z-score with 95% confidence interval (CI) from obtained from Cox Markov proportional hazard model. Table show the possible risk factors for the individual transitions. Z-score was used to standardise variables to a distribution with mean of zero and standard deviation of 1. Abbreviations: see table 1

8.1.4 Goodness of fit assessment

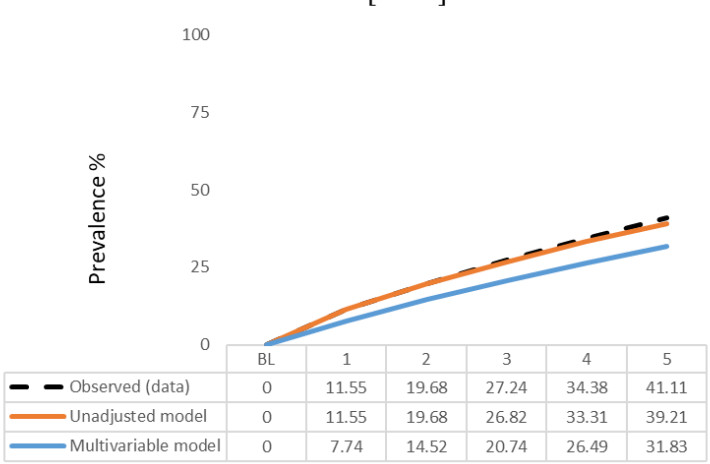
Further evidence for goodness-of-fit was obtained by estimating mean prevalence counts in different states at each time interval. The predicted and observed counts for both the unadjusted and multivariable models are shown in Figure 8.2. The predictions of the unadjusted model closely matched the observed data, whereas the adjusted model overestimated the [*Neither*] and [*Hosp*] states and underestimated [*Dead*].



	BL	1	2	3	4	5
Observed (data)	0	54.10	56.91	52.76	47.75	43.11
Unadjusted model	0	54.39	56.91	52.98	48.46	44.19
Multivariable model	0	58.71	60.38	57.02	53.05	49.24



	BL	1	2	3	4	5
Observed (data)	0	34.36	23.41	20	17.88	15.77
Unadjusted model	0	34.06	23.41	20.2	18.24	16.6
Multivariable model	0	33.55	25.1	22.24	20.46	18.93



	BL	1	2	3	4	5
Observed (data)	0	11.55	19.68	27.24	34.38	41.11
Unadjusted model	0	11.55	19.68	26.82	33.31	39.21
Multivariable model	0	7.74	14.52	20.74	26.49	31.83

Figure 8.2: Goodness-of-fit assessment prediction vs observed

Graphs shows the goodness-of-fit assessment by mean counts in different states at each time interval. The observed (from data) and expected prevalence counts by the fitted models (unadjusted and multivariable) are plotted. Abbreviation: Hosp, Hospital.

8.1.5 Study vignettes

Building other covariates is relatively simple as illustrated by the three examples below.

Additional covariates (i.e., HR, SBP, drugs) were added individually to the multivariable model one by one (Table 8.7). For example, the risk of transitioning from [*BL*] to death is 40% less (HR = 0.60; 95% CI, 0.47 - 0.76) if a patient is on a high-dose beta-blocker (BB) compare to none. Even if the patient was on low doses of beta-blocker, the risk of transitioning to death was still 16% (HR = 0.84; 95% CI, 0.69 – 1.01) less than none. Although these figures are high, their effect on transition rates is marginal (Table 8.8). In contrast, when I included (without/with BB) the heart rate or systolic blood pressure (SBP), the risk of transitioning from any group was flat (HRs approximately 1).

This is confirmed using *-2*log-likelihood* ratio test statistics (LLR), which hardly changed on the introduction of these variables in addition to the fitted model. The base (M1) LLR (28048); with the introduction of multi variables (M2) LLR (27,141), adding BB little change LLR (27,109) and further added HR and SBP even smaller reduction LLR (26,106 and 27,106 respectively) Table 8.8.

Beta Blocker

Transitions	Age-group	Sex	NT-proBNP	NYHA	Haemoglobin	Urea	Albumin	Loop diuretic	Beta blocker	Beta blocker
	≥ 75 years	Male	(ng/L)	III-IV	(g/dL)	(mmol/L)	(g/L)	Yes	Low dose	High dose
BL - Neither	1.07 (0.99, 1.16)	0.9 (0.83, 0.98)	0.78 (0.72, 0.85)	0.79 (0.72, 0.86)	1.04 (1.01, 1.06)	0.99 (0.98, 1)	1.01 (1, 1.02)	0.99 (0.91, 1.08)	1.02 (0.93, 1.12)	1.08 (0.98, 1.19)
BL - Hosp	0.86 (0.77, 0.95)	1.11 (1, 1.23)	1.1 (0.99, 1.21)	1.22 (1.1, 1.35)	0.98 (0.95, 1.01)	0.99 (0.98, 1)	1.01 (1, 1.03)	1.05 (0.94, 1.18)	1.05 (0.93, 1.18)	1.06 (0.93, 1.2)
BL - Dead	1.37 (1.15, 1.64)	1.13 (0.95, 1.35)	2.25 (1.9, 2.66)	1.44 (1.21, 1.7)	0.91 (0.87, 0.96)	1.02 (1.01, 1.03)	0.95 (0.93, 0.97)	1.16 (0.92, 1.45)	**0.84 (0.69, 1.01)**	**0.6 (0.47, 0.76)**
Neither - Neither	0.95 (0.86, 1.05)	0.91 (0.82, 1)	0.92 (0.83, 1.01)	0.94 (0.84, 1.06)	1.03 (0.99, 1.06)	0.99 (0.97, 1)	1 (0.98, 1.01)	0.91 (0.82, 1.01)	1.03 (0.92, 1.15)	1.08 (0.96, 1.21)
Neither - Hosp	1.02 (0.86, 1.22)	1.18 (0.98, 1.41)	1.14 (0.95, 1.37)	1.01 (0.83, 1.23)	0.93 (0.88, 0.98)	1.01 (0.99, 1.03)	1.03 (1, 1.06)	1.23 (1.01, 1.5)	1.05 (0.86, 1.28)	0.89 (0.71, 1.11)
Neither - Dead	1.65 (1.22, 2.23)	1.48 (1.1, 1.99)	1.47 (1.11, 1.95)	1.4 (1.05, 1.86)	1 (0.92, 1.09)	1.03 (1.01, 1.06)	0.94 (0.91, 0.98)	1.74 (1.21, 2.52)	0.72 (0.52, 0.99)	0.73 (0.51, 1.05)
Hosp - Neither	0.91 (0.79, 1.04)	0.85 (0.74, 0.98)	0.89 (0.77, 1.03)	0.81 (0.7, 0.93)	1.04 (0.99, 1.08)	1 (0.98, 1.01)	1.01 (0.99, 1.03)	1.02 (0.88, 1.19)	1.03 (0.88, 1.2)	1.06 (0.89, 1.25)
Hosp - Hosp	0.94 (0.79, 1.11)	1.21 (1.01, 1.45)	1 (0.84, 1.19)	1.19 (1, 1.41)	0.95 (0.9, 1)	0.99 (0.97, 1.01)	1 (0.98, 1.02)	0.94 (0.77, 1.13)	1.01 (0.83, 1.23)	0.99 (0.8, 1.22)
Hosp - Dead	2.17 (1.58, 3)	1.16 (0.85, 1.59)	1.74 (1.26, 2.39)	1.51 (1.13, 2.03)	1.01 (0.92, 1.11)	1.03 (1, 1.05)	0.95 (0.92, 0.99)	1.28 (0.86, 1.89)	0.83 (0.6, 1.16)	0.81 (0.55, 1.19)

Heart rate

Transitions	Age-group	Sex	NT-proBNP	NYHA	Haemoglobin	Urea	Albumin	Loop diuretic	Heart rate
	≥ 75 years	Male	(ng/L)	III-IV	(g/dL)	(mmol/L)	(g/L)	Yes	(bpm)
BL - Neither	1.06 (0.98, 1.15)	0.91 (0.84, 0.99)	0.78 (0.72, 0.85)	0.78 (0.71, 0.86)	1.04 (1.01, 1.06)	0.99 (0.98, 1)	1.01 (1, 1.02)	1 (0.92, 1.09)	1 (1, 1)
BL - Hosp	0.84 (0.76, 0.94)	1.11 (0.99, 1.23)	1.12 (1.01, 1.24)	1.22 (1.1, 1.36)	0.98 (0.95, 1.01)	0.99 (0.98, 1)	1.01 (1, 1.03)	1.07 (0.95, 1.2)	1 (1, 1)
BL - Dead	1.45 (1.21, 1.74)	1.12 (0.94, 1.34)	2.13 (1.79, 2.53)	1.46 (1.23, 1.74)	0.91 (0.86, 0.95)	1.02 (1.01, 1.04)	0.94 (0.92, 0.96)	1.07 (0.86, 1.34)	1 (1, 1.01)
Neither - Neither	0.94 (0.86, 1.04)	0.91 (0.83, 1.01)	0.93 (0.84, 1.03)	0.95 (0.84, 1.06)	1.03 (0.99, 1.06)	0.99 (0.97, 1)	1 (0.99, 1.01)	0.93 (0.84, 1.02)	1 (1, 1)
Neither - Hosp	1.03 (0.86, 1.23)	1.17 (0.98, 1.4)	1.15 (0.96, 1.38)	1.03 (0.84, 1.25)	0.93 (0.87, 0.98)	1.01 (0.99, 1.03)	1.03 (1, 1.05)	1.2 (0.99, 1.46)	1 (0.99, 1)
Neither - Dead	1.75 (1.3, 2.37)	1.44 (1.06, 1.94)	1.31 (0.98, 1.75)	1.35 (1.01, 1.8)	0.99 (0.91, 1.09)	1.04 (1.01, 1.06)	0.94 (0.91, 0.98)	1.65 (1.14, 2.38)	1 (1, 1.01)
Hosp - Neither	0.91 (0.79, 1.05)	0.87 (0.75, 1)	0.89 (0.77, 1.03)	0.8 (0.69, 0.93)	1.03 (0.99, 1.08)	1 (0.98, 1.01)	1.01 (0.99, 1.03)	1.02 (0.88, 1.19)	1 (1, 1.01)
Hosp - Hosp	0.92 (0.77, 1.09)	1.17 (0.98, 1.41)	1.02 (0.86, 1.23)	1.2 (1.01, 1.43)	0.95 (0.9, 1.01)	0.99 (0.97, 1.01)	1 (0.97, 1.02)	0.94 (0.78, 1.14)	1 (0.99, 1)
Hosp - Dead	2.26 (1.63, 3.13)	1.17 (0.85, 1.61)	1.66 (1.21, 2.3)	1.48 (1.1, 2)	1.01 (0.92, 1.1)	1.02 (1, 1.05)	0.95 (0.91, 0.99)	1.23 (0.83, 1.82)	1 (0.99, 1.01)

BP Systolic

Transitions	Age-group	Sex	NT-proBNP (ng/L)	NYHA	Haemoglobin (g/dL)	Urea (mmol/L)	Albumin (g/L)	Loop diuretic	BP Systolic (mmHg)
	≥ 75 years	Male		III-IV				Yes	
BL - Neither	1.05 (0.97, 1.14)	0.91 (0.84, 0.99)	0.79 (0.73, 0.86)	0.79 (0.72, 0.86)	1.04 (1.01, 1.06)	0.99 (0.98, 1)	1.01 (1, 1.02)	1.01 (0.93, 1.1)	1 (1, 1)
BL - Hosp	0.84 (0.76, 0.93)	1.12 (1.01, 1.25)	1.1 (1, 1.22)	1.22 (1.09, 1.35)	0.98 (0.95, 1.01)	0.99 (0.98, 1)	1.01 (1, 1.03)	1.07 (0.96, 1.2)	1 (1, 1)
BL - Dead	1.47 (1.23, 1.75)	1.06 (0.89, 1.26)	2.13 (1.79, 2.52)	1.42 (1.2, 1.69)	0.91 (0.87, 0.96)	1.02 (1.01, 1.04)	0.95 (0.93, 0.96)	1.03 (0.82, 1.29)	0.99 (0.99, 1)
Neither - Neither	0.95 (0.86, 1.05)	0.9 (0.82, 1)	0.92 (0.83, 1.02)	0.94 (0.84, 1.05)	1.03 (0.99, 1.06)	0.99 (0.98, 1)	1 (0.99, 1.01)	0.92 (0.83, 1.01)	1 (1, 1)
Neither - Hosp	1.02 (0.86, 1.22)	1.18 (0.99, 1.42)	1.15 (0.96, 1.38)	1.02 (0.84, 1.24)	0.92 (0.87, 0.98)	1.01 (0.99, 1.03)	1.03 (1, 1.05)	1.24 (1.01, 1.51)	1 (1, 1)
Neither - Dead	1.65 (1.22, 2.22)	1.48 (1.09, 2)	1.5 (1.13, 2)	1.47 (1.11, 1.96)	1 (0.91, 1.09)	1.03 (1.01, 1.06)	0.94 (0.91, 0.98)	1.67 (1.16, 2.41)	1 (1, 1.01)
Hosp - Neither	0.91 (0.79, 1.05)	0.86 (0.74, 0.99)	0.9 (0.78, 1.04)	0.81 (0.7, 0.93)	1.04 (0.99, 1.08)	1 (0.98, 1.01)	1.01 (0.99, 1.03)	1.02 (0.88, 1.19)	1 (1, 1)
Hosp - Hosp	0.93 (0.78, 1.1)	1.21 (1.01, 1.45)	1 (0.84, 1.19)	1.19 (1, 1.41)	0.95 (0.9, 1)	0.99 (0.97, 1.01)	1 (0.98, 1.02)	0.95 (0.79, 1.15)	1 (1, 1)
Hosp - Dead	2.16 (1.57, 2.98)	1.12 (0.82, 1.53)	1.73 (1.26, 2.37)	1.52 (1.13, 2.04)	1.01 (0.92, 1.11)	1.03 (1, 1.05)	0.95 (0.91, 0.98)	1.25 (0.84, 1.84)	1 (1, 1.01)

Table 8.7: Multivariable Cox Markov model with additional covariates to estimated HR

Transitions	Beta-blocker (BB)	Heart rate (HR)	BP Systolic (BPS)
	Transition intensities (95% CI)	**Transition intensities (95% CI)**	**Transition intensities (95% CI)**
BL - Neither	0.59 (0.54, 0.64)	0.59 (0.54, 0.64)	0.59 (0.54, 0.64)
BL - Hosp	0.34 (0.32, 0.35)	0.34 (0.32, 0.35)	0.33 (0.32, 0.35)
BL - Dead	0.08 (0.07, 0.09)	0.08 (0.07, 0.09)	0.08 (0.07, 0.09)
Neither - Neither	0.72 (0.65, 0.79)	0.72 (0.65, 0.79)	0.72 (0.65, 0.79)
Neither - Hosp	0.21 (0.19, 0.23)	0.21 (0.19, 0.23)	0.21 (0.19, 0.23)
Neither - Dead	0.07 (0.06, 0.08)	0.07 (0.06, 0.08)	0.07 (0.06, 0.08)
Hosp - Neither	0.54 (0.5, 0.58)	0.54 (0.51, 0.58)	0.54 (0.51, 0.58)
Hosp - Hosp	0.38 (0.3, 0.46)	0.38 (0.3, 0.45)	0.37 (0.3, 0.45)
Hosp - Dead	0.08 (0.07, 0.1)	0.08 (0.07, 0.1)	0.08 (0.07, 0.1)

Table 8.8: Estimated multivariable TI further adjusting for BB, HR and BPS*Data is expressed as transition intensities with 95% CI (rounded to 2 digits). Adjusted with Age, Sex, NTproBNP, NYHA, Haemoglobin, Urea, Albumin, Loop Diuretic. Abbreviation: See Table 7.1

8.1.6 Sensitivity

I stratified by diagnostic category (phenotypes), as shown in Table 8.9. This can lead to more accurate estimates of transition probabilities and intensities within each subgroup. Healthcare resources can be allocated more efficiently, focusing on those who need the most attention. As discussed earlier in descriptive characteristics, the model has predicted less variation in transition intensities can be seen for HeFREF and HeFPEF ≥ 400. The probability of transitioning to [*Dead*] is lower in HeFPEF 125-399.

Transitions	HeFREF	HeFPEF>=400	HeFPEF>=125-399
	Transition intensities (95% CI)	**Transition intensities (95% CI)**	**Transition intensities (95% CI)**
BL - Neither	0.56 (0.49, 0.63)	0.56 (0.48, 0.63)	0.69 (0.57, 0.81)
BL - Hosp	0.35 (0.33, 0.38)	0.34 (0.32, 0.37)	0.28 (0.25, 0.31)
BL - Dead	0.09 (0.07, 0.1)	0.1 (0.08, 0.12)	0.03 (0.02, 0.05)
Neither - Neither	0.72 (0.61, 0.83)	0.68 (0.56, 0.79)	0.82 (0.67, 0.97)
Neither - Hosp	0.21 (0.18, 0.24)	0.24 (0.21, 0.28)	0.15 (0.12, 0.18)
Neither - Dead	0.07 (0.06, 0.09)	0.08 (0.07, 0.1)	0.03 (0.02, 0.05)
Hosp - Neither	0.52 (0.47, 0.57)	0.53 (0.48, 0.59)	0.59 (0.5, 0.7)
Hosp - Hosp	0.39 (0.28, 0.5)	0.37 (0.25, 0.49)	0.38 (0.18, 0.58)
Hosp - Dead	0.09 (0.07, 0.12)	0.1 (0.07, 0.13)	0.03 (0.01, 0.07)

Table 8.9: Estimated transition intensities stratified by diagnostic categories

8.1.7 Long-term prediction with Sojourn time and Total length of stay

Mean sojourn time[10]: As illustrated in Table 8.10, the maximum mean sojourn time of unadjusted model (M1) and multivariable model (M2) related to [*Neither*] was equal to 3.60 (95% CI: 3.32, 3.84) and 3.44 (95% CI: 3.21, 3.69) cycles respectively. The average amount of time spent in the state [*Hosp*] before making a transition was 1.54 (95% CI: 1.45, 1.64) or 1.61 (95% CI: 1.51, 1.71) cycles. The predicted cycles for both states remain similar without/with covariates. The observed average was 1.8 cycles in the [*Hosp*] state in 5 years. On the other hand, the inclusion of the additional covariate (BB, SBP, HR) in the adjusted model indicates that the sojourn time remains the same. *The covariates appear to influence the transition rates rather than the time spent (cycles) in each state*.

State	**Unadjusted model (M1)**		**Adjusted model (M2)**	
	estimate	SE	estimate	SE
Neither	3.44 (3.21, 3.69)	0.12	3.60 (3.32, 3.84)	0.13
Hosp	1.54 (1.45, 1.64)	0.05	1.61 (1.51, 1.71)	0.05

Table 8.10: Mean sojourn time (in cycles) for the two models (M1 & M2)Abbreviation: SE, standard error; Hosp, Hospitalisation

[10] The mean sojourn time is defined as the average time a patient spends in each state in single stay (which could include consecutive cycles) before making any transition to other states.

The two models can be run through notionally in a theoretical space to the point where all the subjects have died *(*the *total length of stay)*[11]. Table 8.11 shows the projected number of cycles (for both the unadjusted and multivariable model) that an individual will spend in each of the potential states before transitioning to [*Dead*]. Both estimates (that life expectancy is approximately 10 years (7.11 + 2.83) and 13 (9.06 + 3.63) years in the two models) are consistent with the actual observation of 41% mortality at 5 years.

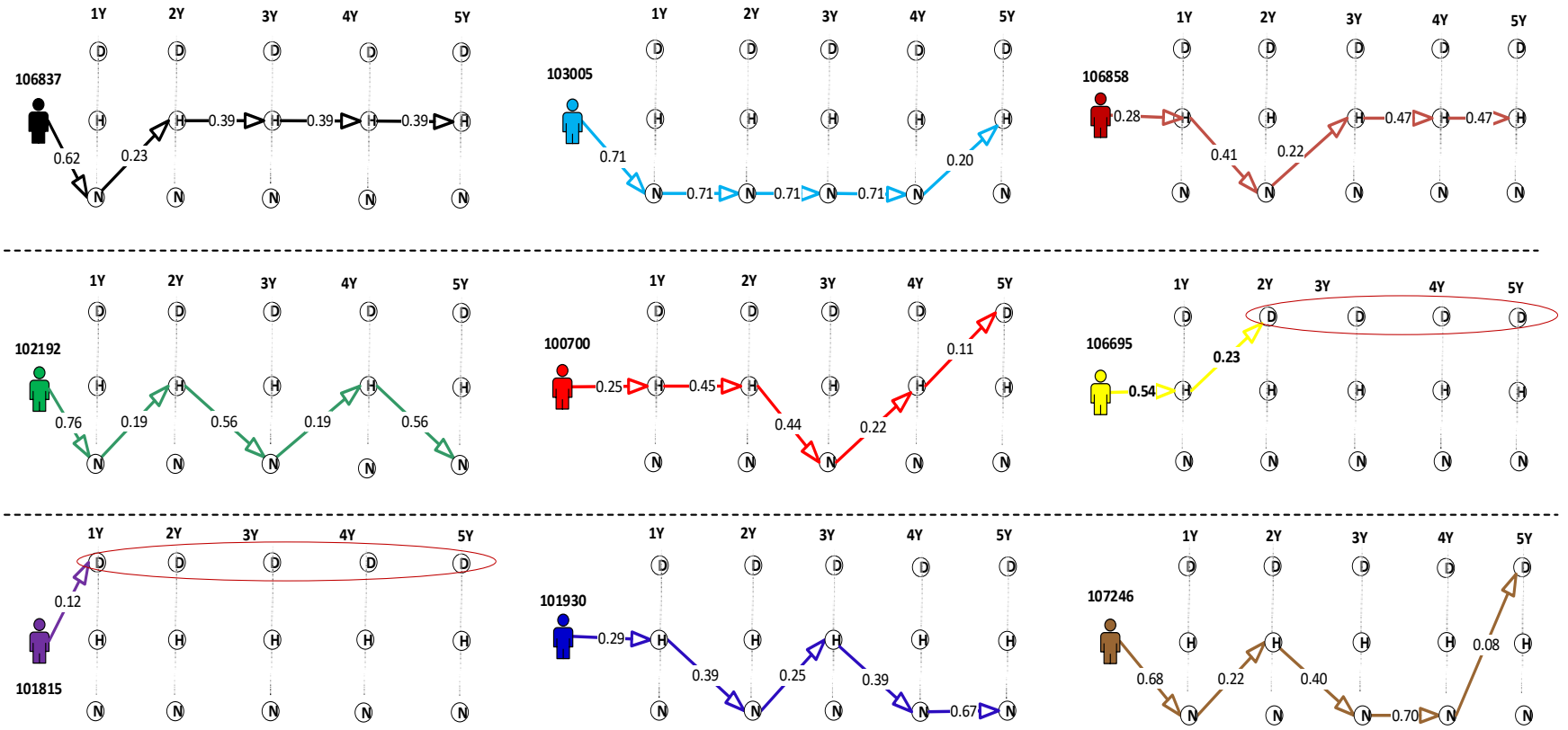
State	Model	
	Unadjusted (M1)	**Adjusted (M2)**
Neither	7.11	9.06
Hosp	2.83	3.63

Table 8.11: Total length of stay (in cycles) for the two models (M1 & M2)

8.1.8 Case trajectories

The same random selection of 9 patients (chapter 2) was made to investigate the prediction made by the adjusted model (Figure 8.3). The individual's prediction (their probability of state transition up to the 5th cycle) based on their covariate values was obtained using equation 7.1. The BL baseline characteristics of these individuals also can be seen. It can be seen that such an analysis would help in determining the causes of unexpected causes for hospitalisations. Note that the model predicted the trajectories based on the baseline values (Table 8.1).

[11] The total length of stay is defined as the anticipated exposure time spent by an individual in each state during the study period before death.



PTNUM	Age_Group (>=75 (1), <75 (0))	Age_AtEvent	Sex (Male(1))	NTproBNP (log10)	NYHA_III_IV	Haemoglobin	Urea	Albumin	Loop_Diuretic_YN	diag	pdiaglu
100700	0	63.4	1	3.12	0	13.90	5.90	36	1	4	HeFREF
101815	1	80.3	0	3.46	0	11.30	23.80	38	0	3	HeFNEF>=400
101930	0	70.3	0	3.39	1	12.10	7.30	39	1	4	HeFREF
102192	0	65.9	0	2.36	0	13.60	4.40	40	1	2	HeFNEF 125-399
103005	1	80.5	0	3.30	0	13.70	5.30	39	1	4	HeFREF
106837	1	84.8	0	4.04	0	11.30	6.40	36	1	4	HeFREF
106858	0	68.3	0	3.08	1	12.50	5.70	37	1	4	HeFREF
107246	1	87.6	0	2.46	1	12.60	4.00	40	1	2	HeFNEF 125-399
106695	1	84.3	1	3.32	1	10.60	10.30	36	1	4	HeFREF

Figure 8.3: Predicted probabilities over time for 9 randomly selected patientsThe table below show the baseline characteristics of these patients and these 8 covariates were used to calculate the probabilities using the equation$qrs = qoexp\,(beta\,z(t))$.

8.2 Discussion

I have used a four-state Markov model to describe the disease trajectory of patients with heart failure. I generated the model by observing events during the first two years of follow up, and the model yielded surprisingly accurate predictions of how a population with heart failure will behave during subsequent years. Trying to understand the trajectory of heart failure in a population of individuals is complicated by the vagaries of life: patients have multiple co-morbidities, and may have events at random that are unrelated to heart failure. Any model attempting to describe progression will thus be limited by the failure of any model to include all possible variables (Kazmi et al. (2022). I attempted to improve the power of the model by including covariates that might have an impact on subsequent outcome, but found that the multivariable model was slightly better predicting cardiovascular (CV) mortality when patients were hospitalised.

When assessing HRs, NT-proBNP (ng/L) was found to be the most influential factor for moving from [*BL*] to [*Hosp*] or to death (i.e., [*Dead*]). Age was the strongest variable for moving to death from being hospitalised or not being hospitalised. It has been noticed that a significant number of patients over the age of 75 years moved from BL to death rather than hospitalisation due to selection bias. Loop diuretic was initially considered as present or absent. E.g. the probability of transitioning from [Hosp] to [Dead] was 0.08 and replacing this variable with Loop dose (mg) change to probability of transitioning 0.10. Loop dose is more sensitive than loop present and absent because it is continuously distributed.

I show that it is *possible* to combine Cox modelling with multistate Markov chains, but that in the instance of HF, it does not enhance the predictive power of a model. However, with the cox modelling I can predict the CV death better. The explanation lies with the fact that the covariates at baseline were included only: subsequent changes to the covariates were ignored by the model. Further work might tackle the complexity of changes in covariates with time, but given the predictive power of the unadjusted model, such an approach might not greatly increase the value of the model. Even when I restricted the analysis to the endpoint of cardiovascular death only, the adjusted model did not perform better than the unadjusted. There is some error in the unadjusted model, but there is no particular pattern of either under or over-prediction.

The study adopts a dual temporal perspective, focusing on both short term (immediate) and long term (extended) prediction. The unadjusted model fits so well with the observed data during subsequent years, it is likely that the analysis of total length of stay is valid. The implication is that with current medical management, a newly assessed patient with heart failure might expect to survive for 10 cycles with approximately three hospitalised cycles in that period. This analysis may be of interest when applied to the population of patients in a clinical trial. Such trials are almost never run to completion (i.e., to the point where all subjects have died), but developing a Markov model by using the transitions during the first two years of a clinical trial (the typical length of cardiovascular trials) to model what the lifetime effect on death and hospitalisations might be, is potentially a helpful way to visualise the impact of an intervention.

The average sojourn time was similar irrespective of whether the covariates were included or not (Table 8.12). In order to further understand the impact of other factors, additional covariates (BB, SBP, HR) were included in the multivariable model. The inclusion had little influence on transition intensities but did not change the sojourn time.

The total length of stay (Table 8.11) also indicating the survival of patients increased by three cycles when clinical features were included in the model. This shows that *the covariates appear to influence the rates of transition and total length of stay rather than the time spent (cycles) in each state.*

An additional value for running such models is to understand the implication for the design of health services. Planning is difficult: often systems are structured in response to the measured incidence of a specific condition without a consideration of the overall impact of a condition might be during the lifetime of an individual patient. Here, I have developed a model that suggests what the implications for the health care system might be given a diagnosis of HF.

In the model, at every cycle, the proportion of patients transitioning from a given state to another remains constant. This finding suggests that the course of heart failure is more linear at a population level than is commonly supposed and, thus, much more predictable if the current state of a patient is known. One consequence of the finding is that that it is the state of the patient at the end of one cycle that carries prognostic information, not the history of preceding states. The past information simply leads to the current state, once you're in that state you don't need to reference past states to predict what will happen next ("memoryless" property).

- **Limitations**

This is an observational study of a large population in a single centre only: the characteristics of the population might not be representative of patients with heart failure as a whole. This is the only hospital providing secondary care for the patients locally, but

it is possible that some patients were admitted to hospital elsewhere. However, deaths are collected centrally and are accurate.

The multivariable model was less accurate than the unadjusted model because it only used covariates from the baseline visit and didn't account for changes in clinical factors over time. The models were not designed to predict the total number of hospital events and/or total stay in hospital (in days) in a particular cycle. The model only considered the possibility of one hospital event within a cycle: if a patient was hospitalised and then died within a single cycle, the hospitalisation was not counted.

The general limitations associated with MSM extend to this work. Fitting an MSM model with covariates can be complicated and computationally intensive due to a potentially large number of potential variables as well as potential parameters: hence, I developed a model with a limited number of covariates and limited it only to those parameters describing the data well.

8.3 Conclusion

A simple 4 state Markov chain model derived from events observed in the first two years of follow up predicts the outcomes of patients with chronic heart failure with a high degree of accuracy. Dynamic risk stratification based on the prediction of an MSM provides greater insight into the clinical trajectory of groups of patients. For this, it requires a rigid categorisation into a finite number of mutually exclusive and exhaustive disease states. The model might be useful in designing health care systems to accommodate the large number of patients with heart failure.

Chapter 9 Conclusion

In this thesis I have successfully used a large dataset of patients being investigated for a possible diagnosis of heart failure to develop dynamic models that effectively describe the disease trajectories for groups of patients as well as individuals. These models are capable of predicting adverse events over a number of consecutive cycles and effectively capturing both short-term and long-term behaviours of the disease. Adopting a dual temporal perspective (focusing on both immediate and extended timeframes) is crucial for understanding how any chronic condition evolves over time.

This thesis provides a range of modelling strategies, each offering a unique perspective on the understanding of risk and dynamic risk modelling of CHF. Depending on the level of detail required, a specific strategy can be employed. Each model, detailed in separate chapters, focused on various states of disease progression. These models have been validated across different time intervals (specially, different cycle durations) to ensure their reliability and accuracy. The study successfully employed machine learning techniques to model disease progression.

The incorporation of clinical covariates into the models has provided deeper insights into how diverse factors influence the disease trajectories of heart failure patients. These models have demonstrated their ability to track changes in health status accurately and predict the risk of death or hospitalisation over time. The methodology developed in this thesis offers a framework that can be adopted to design (and resource) systems tailored for a group of patients with possible heart failure. Importantly, this approach also holds potential for application to other chronic conditions, such as chronic lung or kidney disease, thereby broadening the impact of this work.

9.1 Summary of work and key contributions

9.1.1 Explore the progression of HF using data.

Chapter two provided essential background and context on the progression of chronic heart failure. Here, I highlighted the distinguishing between progressive and non-progressive chronic diseases. Various methodologies and mathematical functions used to model changes health trajectories were described. I highlighted key element required for analysing and modelling changes in patient health status over time. The progression of HF disease was thoroughly explored using extensive Hull LifeLab data in chapter four. Various health states and the risk factors were identified for evaluation of their influence on the disease trajectories and outcomes. I also illustrated the organisation of data into longitudinal health events.

Key contribution: This exploration builds a foundational understanding essential for dynamic risk modelling, which supports further objectives in analysing heart failure progression. Specifically, it provides insights into the various states and transitions patients undergo, informed by observed clinical data and findings detailed in (Kazmi et al., 2022; Kazmi et al., 2024).

9.1.2 Investigate the practical challenges in clinical data

In chapter four I addressed several challenges associated with handling clinical data, including misaligned sample data, missing values and non-normal distributions. I developed detailed strategies to improve the data handling which facilitated the accurate modelling of disease progression in subsequent chapters (5-8). The findings of chapter five laid the ground work for further enhancement in the methodologies, which were elaborated upon in chapter 6. Building on these results, chapter seven introduced the final model.

Key contribution: This iterative process of refinement strengthened the robustness of my data analysis and established a strong foundation for ongoing and future research. The strategies developed for managing clinical data challenges have improved the accuracy and reliability of my predictive models.

9.1.3 Development of dynamic risk models for trajectories in HF

The development of dynamic risk models began with providing patients' baseline characteristics and highlighting the differences between patients with and without OPD follow-up visits. This comparison aimed to identify any unique baseline traits between the two groups, as detailed in **chapter Four**. The data were organised in longitudinal health events format. In **chapter five**, supervised machine learning techniques were applied to classify subsets of the Hull LifeLab population into different classes. The process helped in extracting meaningful insights from the dataset. This also involved identifying if data imbalances could affect model accuracy in later stages. The insight

derived from these analyses was critical for understanding the health trajectories of patients with heart failure.

Key contribution: Building on the groundwork, I developed dynamic risk models using absorbing Markov chains in **chapter seven**. I predicted future outcomes based on earlier events in a patient's history, and analysed the influence of age and sex on disease progression. I achieved the study's aim by demonstrating that dynamic risk stratification provides a unique perspective compared to traditional models of disease progression. This approach views the clinical trajectory of patients collectively while also allowing for individual trajectory predictions, enabling further survival analysis. I further expanded dynamic risk model by incorporating clinical covariates to examine their impact on transitions between clinical states in order to broaden the prediction scale to cover both short-term and long-term behaviours.

9.1.4 Linearity in Heart Failure progression:

The analysis in chapter six and seven suggests that at every cycle, the proportion of patients transitioning from one state to another remains constant. One consequence of the finding is that the patient's state at the end of one cycle carries prognostic information, not the history of preceding states. The history is "hidden" in the present state.

Key contribution: The findings suggest that the course of heart failure is representative of a linear model at a population level than is commonly supposed, and, thus, much more predictable for an individual if the current state of a patient is known.

The structure of the thesis and the objectives provides an understanding and application of a multistate risk modelling framework to the problem of progressive chronic heart failure (CHF). This comprehensive approach has ensured that each stage of the modelling process is robust and that the conclusions drawn are well-supported by empirical evidence.

9.2 Limitations and future directions

This is an observational study of a large population in a single centre only: the characteristics of the population might not be representative of patients with heart failure as a whole. The data are taken from the only hospital providing secondary care for the patients locally, but some patients may have been admitted to the hospital elsewhere. However, deaths are collected centrally and are accurate.

The multivariable model's predictions were less accurate than the unadjusted model reflects the fact that I was limited to the covariates collected at the baseline visit and did not consider changes in clinical covariates from baseline. The models were not designed to predict the total number of hospital events and/or total stay in hospital (in days) in a particular cycle. The model only considered the possibility of one hospital event within a cycle: if a patient was hospitalised and died within a single cycle, the hospitalisation was not counted.

The general limitations associated with MSM extend to this work. Fitting an MSM model with covariates can be complicated and computationally intensive due to the potentially large number of potential variables and potential parameters: hence, I developed a model with a limited number of covariates and limited it only to those parameters describing the data well.

Recommendations for future work include further validation of the models I have developed in different populations and settings and exploring the integration of newer data sources and emerging technologies in model refinement and implementation.

Reference list / Bibliography

Reference List

Abel, A. A. I., Samuel, N. A., Cuthbert, J. J., Brown, O. I., Pellicori, P., Kazmi, S., Cleland, J. G. F., Johnson, M. J. & Clark, A. L. (2024) Hospital admissions in the last year of life of patients with heart failure. *Eur Heart J Qual Care Clin Outcomes*, 10(2), 168-175.

Alejandro, C. & Banco, C. (2012) Constructing a Credit Risk Scorecard using Predictive Clusters. *SAS Global Forum 2012*, 128-2012, 1-13.

Allen, L. A., Stevenson, L. W., Grady, K. L., Goldstein, N. E., Matlock, D. D., Arnold, R. M., Cook, N. R., Felker, G. M., Francis, G. S., Hauptman, P. J., Havranek, E. P., Krumholz, H. M., Mancini, D., Riegel, B., Spertus, J. A., American Heart, A., Council on Quality of, C., Outcomes, R., Council on Cardiovascular, N., Council on Clinical, C., Council on Cardiovascular, R., Intervention, Council on Cardiovascular, S. & Anesthesia (2012) Decision making in advanced heart failure: a scientific statement from the American Heart Association. *Circulation*, 125(15), 1928-52.

Andersen, P. K. & Keiding, N. (2002) Multi-state models for event history analysis. *Stat Methods Med Res*, 11(2), 91-115.

Anderson, B. (2014) Probability and the changing shape of response distributions for orientation. *J Vis*, 14(13), 15.

Anderson, B. & Druker, M. (2013) Attention improves perceptual quality. *Psychon Bull Rev*, 20(1), 120-7.

Benkel, I., Arnby, M. & Molander, U. (2020) Living with a chronic disease: A quantitative study of the views of patients with a chronic disease on the change in their life situation. *SAGE Open Med*, 8, 2050312120910350.

Bland, J. M. & Altman, D. G. (1996) Statistics notes. Logarithms. *BMJ*, 312(7032), 700.

Bohacik, J., Kambhampati, C., Davis, D. N. & Cleland, J. (2013) Alternating decision tree applied to risk assessment of heart failure patients. *Journal of Information Technologies*, 6, 25-33.

Bolman, R. M. & Black, S. M. (2003) Open cardiac repair under direct vision: F. John Lewis and the University of Minnesota. *Journal of Cardiac Surgery*, 18(4), 328-332.

Braunwald, E. (2015) The war against heart failure: the Lancet lecture. *Lancet*, 385(9970), 812-24.

Brock, G. N., Shaffer, J. R., Blakesley, R. E., Lotz, M. J. & Tseng, G. C. (2008) Which missing value imputation method to use in expression profiles: a comparative study and two selection schemes. *BMC Bioinformatics*, 9, 12.

Buda, M., Maki, A. & Mazurowski, M. A. (2018) A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw*, 106, 249-259.

Campbell, D. & Stanley, J. (1963) Experimental and quasi-experimental designs for research. *Chicago, IL: Rand McNally;*.

Castaneda, J. & Gerritse, B. (2010) Appraisal of Several Methods to Model Time to Multiple Events per Subject: Modelling Time to Hospitalizations and Death. *Revista Colombiana De Estadistica*, 33(1), 43-61.

Chan, D. C., Heidenreich, P. A., Weinstein, M. C. & Fonarow, G. C. (2008) Heart failure disease management programs: a cost-effectiveness analysis. *Am Heart J*, 155(2), 332-8.

Chang, W. C., Kaul, P., Fu, Y., Westerhout, C. M., Granger, C. B., Mahaffey, K. W., Wallentin, L., Van de Werf, F., Armstrong, P. W. & Investigators, A.-. (2006) Forecasting mortality: dynamic assessment of risk in ST-segment elevation acute myocardial infarction. *Eur Heart J*, 27(4), 419-26.

Cismondi, F., Fialho, A. S., Vieira, S. M., Reti, S. R., Sousa, J. M. & Finkelstein, S. N. (2013) Missing data in medical databases: impute, delete or classify? *Artif Intell Med*, 58(1), 63-72.

Cleland, J. G., Henry, D., Hardman, S., McDonagh, T., National Heart Failure Audit Team for, E. & Wales (2012-13) National Heart Failure Audit 2012-13, 87.

Cleland, J. G., McDonagh, T., Rigby, A. S., Yassin, A., Whittaker, T., Dargie, H. J., National Heart Failure Audit Team for, E. & Wales (2011) The national heart failure audit for England and Wales 2008-2009. *Heart*, 97(11), 876-86.

Cohen, B., Vawdrey, D. K., Liu, J., Caplan, D., Furuya, E. Y., Mis, F. W. & Larson, E. (2015) Challenges Associated With Using Large Data Sets for Quality Assessment and Research in Clinical Settings. *Policy Polit Nurs Pract*, 16(3-4), 117-24.

Collins, L. M. (2006) Analysis of longitudinal data: the integration of theoretical model, temporal design, and statistical model. *Annu Rev Psychol*, 57, 505-28.

Collins, L. M. & Horn, J. L. (1991) *Best methods for the analysis of change: Recent advances, unanswered questions, future directions*. Washington, DC, US: American Psychological Association.

Colombo, G. L., Caruggi, M., Ottolini, C. & Maggioni, A. P. (2008) Candesartan in heart failure: assessment of reduction in mortality and morbidity (CHARM) and resource utilization and costs in Italy. *Vasc Health Risk Manag*, 4(1), 223-34.

Comstock Barker, P. & Scherer, J. S. (2017) Illness Trajectories: Description and Clinical Use #326. *J Palliat Med*, 20(4), 426-427.

Corbin, J. M. (1998) The Corbin and Strauss Chronic Illness Trajectory model: an update. *Sch Inq Nurs Pract*, 12(1), 33-41.

Counsell, N., Cortina-Borja, M., Lehtonen, A. & Stein, A. (2011) Modelling psychiatric measures using Skew-Normal distributions. *Eur Psychiatry*, 26(2), 112-4.

Cudeck, R. & Klebe, K. J. (2002) Multiphase mixed-effects models for repeated measures data. *Psychol Methods*, 7(1), 41-63.

Cuthbert, J. J., Brown, O. I., Pellicori, P., Dobbs, K., Bulemfu, J., Kazmi, S., Sokoreli, I., Pauws, S. C., Riistama, J. M., Cleland, J. G. F. & Clark, A. L. (2024) Medicines optimization prior to discharge in patients admitted to hospital with heart failure. *ESC Heart Fail*, 11(2), 950-961.

David, P. D. (2011) Measuring Skewness: A Forgotten Statistic? *Journal of Statistics Education*, 19.

Diaz, R., Behr, J., Kumar, S. & Britton, B. (2015) Modeling Chronic Disease Patient Flows Diverted from Emergency Departments to Patient-Centered Medical Homes. *IIE Trans Healthc Syst Eng*, 5(4), 268-285.

Fairclough, D. L. (2010) *Design and Analysis of Quality of Life Studies in Clinical Trials*.CRC Press.

Frohlich, H., Rosenfeld, N., Tager, T., Goode, K., Kazmi, S., Hole, T., Katus, H. A., Atar, D., Cleland, J. G. F., Agewall, S., Clark, A. L., Frankenstein, L. & Grundtvig, M. (2019) Epidemiology and long-term outcome in outpatients with chronic heart failure in Northwestern Europe. *Heart*, 105(16), 1252-1259.

Futoma, J., Morris, J. & Lucas, J. (2015) A comparison of models for predicting early hospital readmissions. *J Biomed Inform*, 56, 229-38.

Glasgow Caledonian University (2012) Probability and Probability Distributions. *SCHOOL OF ENGINEERING & BUILT ENVIRONMENT*.

Goel, M. K., Khanna, P. & Kishore, J. (2010) Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res*, 1(4), 274-8.

Gruger, J., Kay, R. & Schumacher, M. (1991) The validity of inferences based on incomplete observations in disease state models. *Biometrics*, 47(2), 595-605.

Gupta, A., Kumar, L., Jain, R. & Nagrath, P. (2020) Heart Disease Prediction Using Classification (Naive Bayes), *Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019)*. Singapore, 2020//. Springer Singapore.

Harris, C. W. & Youth, S. S. R. C. C. o. P. D. i. (1967) *Problems in Measuring Change, Edited by Chester W. Harris. Proceedings of a Conference Sponsored by the Committee on Personality Development in Youth of the Social Science Research Council, 1962*.

Henly, S. J., Kallas, K. D., Klatt, C. M. & Swenson, K. K. (2003) The notion of time in symptom experiences. *Nurs Res*, 52(6), 410-7.

Henly, S. J., Wyman, J. F. & Findorff, M. J. (2011) Health and illness over time: the trajectory perspective in nursing science. *Nurs Res*, 60(3 Suppl), S5-14.

Hupcey, J. E., Penrod, J. & Fenstermacher, K. (2009) A Model of Palliative Care for Heart Failure. *American Journal of Hospice & Palliative Medicine*, 26(5), 399-404.

Hutchinson, K., Pellicori, P., Dierckx, R., Cleland, J. G. & Clark, A. L. (2014) Remote telemonitoring for patients with heart failure: might monitoring pulmonary artery pressure become routine? *Expert Rev Cardiovasc Ther*, 12(8), 1025-33.

Ieva, F., Jackson, C. H. & Sharples, L. D. (2017) Multi-state modelling of repeated hospitalisation and death in patients with heart failure: The use of large administrative databases in clinical epidemiology. *Stat Methods Med Res*, 26(3), 1350-1372.

Ingle, L., Cleland, J. G. & Clark, A. L. (2014) The long-term prognostic significance of 6-minute walk test distance in patients with chronic heart failure. *Biomed Res Int*, 2014, 505969.

Inglis, S. C., Clark, R. A., McAlister, F. A., Ball, J., Lewinter, C., Cullington, D., Stewart, S. & Cleland, J. G. (2010) Structured telephone support or telemonitoring programmes for patients with chronic heart failure. *Cochrane Database Syst Rev*(8), CD007228.

Jackson, C. (2019) Multi-state modelling with R: the msm package.

Jiang, W., Siddiqui, S., Barnes, S., Barouch, L. A., Korley, F., Martinez, D. A., Toerper, M., Cabral, S., Hamrock, E. & Levin, S. (2019) Readmission Risk Trajectories for Patients With Heart Failure Using a Dynamic Prediction Approach: Retrospective Study. *JMIR Med Inform*, 7(4), e14756.

Kazmi, S., Kambhampati, C., Cleland, J. G. F., Cuthbert, J., Kazmi, K. S., Pellicori, P., Rigby, A. S. & Clark, A. L. (2022) Dynamic risk stratification using Markov chain modelling in patients with chronic heart failure. *Esc Heart Failure*.

Kazmi, S., Kambhampati, C., Rigby, A. S., Cleland, J. G. F., Kazmi, K. S., Cuthbert, J., Pellicori, P. & Clark, A. L. (2024) Disease progression in chronic heart failure is linear: Insights from multistate modelling. *European Journal of Heart Failure*.

Kesavaraj, G. & Sukumaran, S. (2013) A Study On Classification Techniques in Data Mining. *2013 Fourth International Conference on Computing, Communications and Networking Technologies (Icccnt)*.

Ketchum, E. S. & Levy, W. C. (2011) Establishing prognosis in heart failure: a multimarker approach. *Prog Cardiovasc Dis*, 54(2), 86-96.

Khand, A. U., Gemmell, I., Rankin, A. C. & Cleland, J. G. (2001) Clinical events leading to the progression of heart failure: insights from a national database of hospital discharges. *Eur Heart J*, 22(2), 153-64.

Kheirbek, R. E., Alemi, F., Citron, B. A., Afaq, M. A., Wu, H. & Fletcher, R. D. (2013) Trajectory of illness for patients with congestive heart failure. *J Palliat Med*, 16(5), 478-84.

Koulaouzidis, G., Barrett, D., Mohee, K. & Clark, A. L. (2019) Telemonitoring in subjects with newly diagnosed heart failure with reduced ejection fraction: From clinical research to everyday practice. *J Telemed Telecare*, 25(3), 167-171.

Krajewska, J., Chmielik, E. & Jarzab, B. (2017) Dynamic risk stratification in the follow-up of thyroid cancer: what is still to be discovered in 2017? *Endocr Relat Cancer*, 24(11), R387-R402.

Le-Rademacher, J. G., Therneau, T. M. & Ou, F. S. (2022) The Utility of Multistate Models: A Flexible Framework for Time-to-Event Data. *Curr Epidemiol Rep*, 9(3), 183-189.

Levin, S., Toerper, M., Hamrock, E., Hinson, J. S., Barnes, S., Gardner, H., Dugas, A., Linton, B., Kirsch, T. & Kelen, G. (2018) Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index. *Ann Emerg Med*, 71(5), 565-574 e2.

Li Peng et al (2015) Missing Value Imputation Method Based on Density Clustering and Grey Relational Analysis. *Science & Engineering Research Support Society*, 10(11), 133-142.

Lubkin, I. M. & Larsen, P. D. (2013) *Chronic illness: impact and interventions*. Translated from English by, 8th edition. Burlington, Mass: Jones & Bartlett Learning.

Lyons, J., Akbari, A., Abrams, K. R., Azcoaga Lorenzo, A., Ba Dhafari, T., Chess, J., Denaxas, S., Fry, R., Gale, C. P., Gallacher, J., Griffiths, L. J., Guthrie, B., Hall, M., Jalali-Najafabadi, F., John, A., MacRae, C., McCowan, C., Peek, N., O'Reilly, D., Rafferty, J., Lyons, R. A. & Owen, R. K. (2023) Trajectories in chronic disease accrual and mortality across the lifespan in Wales, UK (2005-2019), by area deprivation profile:

linked electronic health records cohort study on 965,905 individuals. *Lancet Reg Health Eur*, 32, 100687.

Ma, J., Chan, W., Tsai, C. L., Xiong, M. & Tilley, B. C. (2015) Analysis of transtheoretical model of health behavioral changes in a nutrition intervention study - a continuous time Markov chain model with Bayesian approach. *Stat Med*, 34(27), 3577-89.

Maheswari, S. & Pitchai, R. (2019) Heart Disease Prediction System Using Decision Tree and Naive Bayes Algorithm. *Curr Med Imaging Rev*, 15(8), 712-717.

Masini, G., Graham, F. J., Pellicori, P., Cleland, J. G. F., Cuthbert, J. J., Kazmi, S., Inciardi, R. M. & Clark, A. L. (2022) Criteria for Iron Deficiency in Patients With Heart Failure. *J Am Coll Cardiol*, 79(4), 341-351.

McGrath, J. E. & Tschan, F. (2004) *Temporal matters in social psychology: Examining the role of time in the lives of groups and individuals*. Washington, DC, US: American Psychological Association.

McIlvennan, C. K. & Allen, L. A. (2016) Palliative care in patients with heart failure. *BMJ*, 353, i1010.

McMurray, J. J., Adamopoulos, S., Anker, S. D., Auricchio, A., Bohm, M., Dickstein, K., Falk, V., Filippatos, G., Fonseca, C., Gomez-Sanchez, M. A., Jaarsma, T., Kober, L., Lip, G. Y., Maggioni, A. P., Parkhomenko, A., Pieske, B. M., Popescu, B. A., Ronnevik, P. K., Rutten, F. H., Schwitter, J., Seferovic, P., Stepinska, J., Trindade, P. T., Voors, A. A., Zannad, F., Zeiher, A., Task Force for the, D., Treatment of, A., Chronic Heart Failure of the European Society of, C., Bax, J. J., Baumgartner, H., Ceconi, C., Dean, V., Deaton, C., Fagard, R., Funck-Brentano, C., Hasdai, D., Hoes, A., Kirchhof, P., Knuuti, J., Kolh, P., McDonagh, T., Moulin, C., Popescu, B. A., Reiner, Z., Sechtem, U., Sirnes, P. A., Tendera, M., Torbicki, A., Vahanian, A., Windecker, S., McDonagh, T., Sechtem, U., Bonet, L. A., Avraamides, P., Ben Lamin, H. A., Brignole, M., Coca, A., Cowburn, P., Dargie, H., Elliott, P., Flachskampf, F. A., Guida, G. F., Hardman, S., Iung, B., Merkely, B., Mueller, C., Nanas, J. N., Nielsen, O. W., Orn, S., Parissis, J. T., Ponikowski, P. & Guidelines, E. S. C. C. f. P. (2012) ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail*, 14(8), 803-69.

Megari, K. (2013) Quality of Life in Chronic Disease Patients. *Health Psychol Res*, 1(3), e27.

Meira-Machado, L., Cadarso-Suarez, C. & de Una-Alvarez, J. (2007) tdc.msm: an R library for the analysis of multi-state survival data. *Comput Methods Programs Biomed*, 86(2), 131-40.

Meira-Machado, L., de Una-Alvarez, J., Cadarso-Suarez, C. & Andersen, P. K. (2009) Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res*, 18(2), 195-222.

Mhoon, K. B., Chan, W., Del Junco, D. J. & Vernon, S. W. (2010) A Continuous-Time Markov Chain Approach Analyzing the Stages of Change Construct from a Health Promotion Intervention. *JP J Biostat*, 4(3), 213-226.

Moshkovich, O., Benjamin, K., Hall, K., Murphy, R., von Maltzahn, R., Gorsh, B., Sikirica, V., Saini, R. & Sprecher, D. (2020) Correction to: Development of a conceptual model and patient-reported outcome measures for assessing symptoms and functioning in patients with heart failure. *Qual Life Res*, 29(10), 2849.

National Heart Failure Audit Report UK (2019) *National Heart Failure Audit*. UK: Available online: https://www.nicor.org.uk/wp-content/uploads/2019/09/Heart-Failure-2019-Report-final.pdf [Accessed.

National Heart Failure Audit Team for, E. (2010) Chronic Heart Failure: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care, *Chronic Heart Failure: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care: Partial Update*. Translated from eng by. National Institute for Health and Clinical Excellence: Guidance. London, 1-49.

Naumzik, C., Feuerriegel, S. & Nielsen, A. M. (2023) Data-driven dynamic treatment planning for chronic diseases. *European Journal of Operational Research*, 305(2), 853-867.

Neale, B. (2015) Time and the lifecourse: perspectives from qualitative longitudinal research, in Nancy, W. & Irene, H. (eds), *Researching the Lifecourse*. Bristol, UK: Policy Press, 25-42.

Nesselroade, J. & Ram, N. (2004) Studying Intraindividual Variability: What We Have Learned That Will Help Us Understand Lives in Context. *Research in Human Development*, 1, 9-29.

Nguena Nguefack, H. L., Page, M. G., Katz, J., Choiniere, M., Vanasse, A., Dorais, M., Samb, O. M. & Lacasse, A. (2020) Trajectory Modelling Techniques Useful to Epidemiological Research: A Comparative Narrative Review of Approaches. *Clin Epidemiol*, 12, 1205-1222.

NICE (2010) Chronic Heart Failure: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care, *Chronic Heart Failure: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care: Partial Update*. Translated from eng by. National Institute for Health and Clinical Excellence: Guidance. London, 1-49.

NICE (2018a) *Chronic heart failure in adults: Diagnosis and management*. Available online: https://www.nice.org.uk/guidance/ng106/evidence/full-guideline-pdf-6538850029 [Accessed.

NICE (2018b) Diagnosis and management of adults with chronic heart failure: summary of updated NICE guidance, *BMJ*, 2018/09/28 edition, k4080.

Nikolaidou, T., Pellicori, P., Zhang, J., Kazmi, S., Goode, K. M., Cleland, J. G. & Clark, A. L. (2018) Prevalence, predictors, and prognostic implications of PR interval prolongation in patients with heart failure. *Clin Res Cardiol*, 107(2), 108-119.

Ohu, I., Benny, P. K., Rodrigues, S. & Carlson, J. N. (2020) Applications of machine learning in acute care research. *J Am Coll Emerg Physicians Open*, 1(5), 766-772.

Olivier, J., Johnson, W. D. & Marshall, G. D. (2008) The logarithmic transformation and the geometric mean in reporting experimental IgE results: what are they and when and why to use them? *Ann Allergy Asthma Immunol*, 100(4), 333-7.

Pearce, N. (1996) Traditional epidemiology, modern epidemiology, and public health. *Am J Public Health*, 86(5), 678-83.

Pocock, S. J., Ariti, C. A., McMurray, J. J., Maggioni, A., Kober, L., Squire, I. B., Swedberg, K., Dobson, J., Poppe, K. K., Whalley, G. A., Doughty, R. N. & Meta-Analysis Global Group in Chronic Heart, F. (2013) Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J*, 34(19), 1404-13.

Ponikowski, P., Voors, A. A., Anker, S. D., Bueno, H., Cleland, J. G., Coats, A. J., Falk, V., Gonzalez-Juanatey, J. R., Harjola, V. P., Jankowska, E. A., Jessup, M., Linde, C., Nihoyannopoulos, P., Parissis, J. T., Pieske, B., Riley, J. P., Rosano, G. M., Ruilope, L. M., Ruschitzka, F., Rutten, F. H., van der Meer, P., Authors/Task Force, M. &

Document, R. (2016) 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail*, 18(8), 891-975.

Poolsawad, N., Moore, L., Kambhampati, C. & Cleland, J. (2012) Handling missing values in data mining - A case study of heart failure dataset. *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012)*.

Poolsawad, N., Moore, L., Kambhampati, C. & Cleland, J. G. F. (2014) Issues in the Mining of Heart Failure Datasets. *International Journal of Automation and Computing*, 11(2), 162-179.

Putter, H., Fiocco, M. & Geskus, R. B. (2007) Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*, 26(11), 2389-430.

Rigby, A. S. (1991) Development of a Scoring System to Assist in the Diagnosis of Rheumatoid Arthritis. *Methods Inf Med*, 30(01), 23-29.

Roger, V. L. (2013) Epidemiology of heart failure. *Circ Res*, 113(6), 646-59.

Sai Krishna Reddy, V., Meghana, P., Subba Reddy, N. V. & Ashwath Rao, B. (2022) Prediction on Cardiovascular disease using Decision tree and Naïve Bayes classifiers. *Journal of Physics: Conference Series*, 2161(1), 012015.

Saito, T. & Rehmsmeier, M. (2015) The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *Plos One*, 10(3).

Salonen, J., Tikanmäki, H. & Nummi, T. (2019) Using trajectory analysis to test and illustrate microsimulation outcomes. *IJM*, 12(2), 3-17.

Sargent, M., Parker, E., Syme, J. & Kilambi, V. (2024) Developing Models and Metrics to Assess the Impacts of Complexity in Operational Settings.

Sato, R. C. & Zouain, D. M. (2010) Markov Models in health care. *Einstein (Sao Paulo)*, 8(3), 376-9.

Schröger, E., Roeber, U. & Coy, N. (2023) Markov chains as a proxy for the predictive memory representations underlying mismatch negativity. *Frontiers in Human Neuroscience*, 17.

Shih, H. C., Chou, P., Liu, C. M. & Tung, T. H. (2009) Retraction: estimation of progression of multi-state chronic disease using the Markov model and prevalence pool concept. *BMC Med Inform Decis Mak*, 9, 45.

Shiraishi, Y., Nagai, T., Kohsaka, S., Goda, A., Nagatomo, Y., Mizuno, A., Kohno, T., Rigby, A., Fukuda, K., Yoshikawa, T., Clark, A. L. & Cleland, J. G. F. (2018) Outcome of hospitalised heart failure in Japan and the United Kingdom stratified by plasma N-terminal pro-B-type natriuretic peptide. *Clin Res Cardiol*, 107(12), 1103-1110.

Shoaib, A., Farag, M., Nasir, M., John, J., Gupta, S., Pellicori, P., Antony, R., Perveen, R., Rigby, A., Goode, K. M., Yassin, A., Clark, A. L. & Cleland, J. G. (2016) Is the diagnostic coding position of acute heart failure related to mortality? A report from the Euro Heart Failure Survey-1. *Eur J Heart Fail*, 18(5), 556-63.

Shoaib, A., Mamas, M. A., Ahmad, Q. S., McDonagh, T. M., Hardman, S. M. C., Rashid, M., Butler, R., Duckett, S., Satchithananda, D., Nolan, J., Dargie, H. J., Clark, A. L. & Cleland, J. G. F. (2019) Characteristics and outcome of acute heart failure patients according to the severity of peripheral oedema. *Int J Cardiol*, 285, 40-46.

Singer, J. D. & Willett, J. B. (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*.Oxford University Press, USA.

Singer, P., Helic, D., Taraghi, B. & Strohmaier, M. (2014) Detecting memory and structure in human navigation patterns using Markov chain models of varying order. *PLoS One*, 9(7), e102070.

Skrondal, A. & Rabe-Hesketh, S. (2004) *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*.CRC Press.

Sokoreli, I., de Vries, J. J. G., Riistama, J. M., Pauws, S. C., Steyerberg, E. W., Tesanovic, A., Geleijnse, G., Goode, K. M., Crundall-Goode, A., Kazmi, S., Cleland, J. G. & Clark, A. L. (2016) Depression as an independent prognostic factor for all-cause mortality after a hospital admission for worsening heart failure. *International Journal of Cardiology*, 220, 202-207.

Sokoreli, I., Pauws, S. C., Steyerberg, E. W., de Vries, G. J., Riistama, J. M., Tesanovic, A., Kazmi, S., Pellicori, P., Cleland, J. G. & Clark, A. L. (2018) Prognostic value of psychosocial factors for first and recurrent hospitalizations and mortality in heart failure patients: insights from the OPERA-HF study. *Eur J Heart Fail*, 20(4), 689-696.

Sonnenberg, F. A. & Beck, J. R. (1993) Markov models in medical decision making: a practical guide. *Med Decis Making*, 13(4), 322-38.

Sutradhar, R. & Barbera, L. (2014) A Markov multistate analysis of the relationship between performance status and death among an ambulatory population of cancer patients. *Palliat Med*, 28(2), 184-90.

Sutradhar, R., Barbera, L., Seow, H., Howell, D., Husain, A. & Dudgeon, D. (2011) Multistate analysis of interval-censored longitudinal data: application to a cohort study on performance status among patients diagnosed with cancer. *Am J Epidemiol*, 173(4), 468-75.

Taylor, C. J., Ordonez-Mena, J. M., Roalfe, A. K., Lay-Flurrie, S., Jones, N. R., Marshall, T. & Hobbs, F. D. R. (2019) Trends in survival after a diagnosis of heart failure in the United Kingdom 2000-2017: population based cohort study. *BMJ*, 364, l223.

Uhry, Z., Hedelin, G., Colonna, M., Asselain, B., Arveux, P., Rogel, A., Exbrayat, C., Guldenfels, C., Courtial, I., Soler-Michel, P., Molinie, F., Eilstein, D. & Duffy, S. W. (2010) Multi-state Markov models in cancer screening evaluation: a brief review and case study. *Stat Methods Med Res*, 19(5), 463-86.

Upshaw, J. N., Konstam, M. A., Klaveren, D. v., Noubary, F., Huggins, G. S. & Kent, D. M. (2016) Multistate Model to Predict Heart Failure Hospitalizations and All-Cause Mortality in Outpatients With Heart Failure With Reduced Ejection Fraction. *Circulation: Heart Failure*, 9(8), e003146.

Wasikowski, M. & Chen, X. W. (2010) Combating the Small Sample Class Imbalance Problem Using Feature Selection. *Ieee Transactions on Knowledge and Data Engineering*, 22(10), 1388-1400.

Wolfson, J., Bandyopadhyay, S., Elidrisi, M., Vazquez-Benitez, G., Vock, D. M., Musgrove, D., Adomavicius, G., Johnson, P. E. & O'Connor, P. J. (2015) A Naive Bayes machine learning approach to risk prediction using censored, time-to-event data. *Stat Med*, 34(21), 2941-57.

Writing Group, M., Lloyd-Jones, D., Adams, R. J., Brown, T. M., Carnethon, M., Dai, S., De Simone, G., Ferguson, T. B., Ford, E., Furie, K., Gillespie, C., Go, A., Greenlund, K., Haase, N., Hailpern, S., Ho, P. M., Howard, V., Kissela, B., Kittner, S., Lackland, D., Lisabeth, L., Marelli, A., McDermott, M. M., Meigs, J., Mozaffarian, D., Mussolino, M., Nichol, G., Roger, V. L., Rosamond, W., Sacco, R., Sorlie, P., Roger, V. L., Thom, T., Wasserthiel-Smoller, S., Wong, N. D., Wylie-Rosett, J., American Heart Association Statistics, C. & Stroke Statistics, S. (2010) Heart disease and stroke statistics--2010 update: a report from the American Heart Association. *Circulation*, 121(7), e46-e215.

Wu, S. J. & Chu, M. T. (2017) Markov chains with memory, tensor formulation, and the dynamics of power iteration. *Applied Mathematics and Computation*, 303, 226-239.

Yashinski, A. (2021) Absorbing Markov Chains.

Zhang, J., Goode, K. M., Rigby, A., Balk, A. H. M. M. & Cleland, J. G. (2013) Identifying patients at risk of death or hospitalisation due to worsening heart failure using decision tree analysis: evidence from the Trans-European Network-Home-Care Management System (TEN-HMS) study. *International journal of cardiology*, 163(2), 149-56.

Zhang, J., Pellicori, P., Pan, D., Dierckx, R., Clark, A. L. & Cleland, J. G. F. (2018a) Dynamic risk stratification using serial measurements of plasma concentrations of natriuretic peptides in patients with heart failure. *Int J Cardiol*, 269, 196-200.

Zhang, J., Pellicori, P., Pan, D., Dierckx, R., Clark, A. L. & Cleland, J. G. F. (2018b) Dynamic risk stratification using serial measurements of plasma concentrations of natriuretic peptides in patients with heart failure. *International Journal of Cardiology*, 269, 196-200.

Zhang, Y., Kambhampati, C., Davis, D. N., Goode, K. & Cleland, J. (2012) A comparative study of missing value imputation with multiclass classification for clinical heart failure data. *Proceedings - 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2012*.

Zupan, B., Demsar, J., Kattan, M. W., Beck, J. R. & Bratko, I. (2000) Machine learning for survival analysis: a case study on recurrence of prostate cancer. *Artif Intell Med*, 20(1), 59-75.