University of Cape Town

STA5000W

MASTER OF SCIENCE IN MATHEMATICAL STATISTICS

Modelling multivariate longitudinal outcomes and time-to-event data

Author:
Modiehi Theletsane

Supervisor:
Associate Professor Freedom Gumedze

A thesis submitted to the Department of Statistical Sciences in fulfilment of the requirements for the degree Master of Science in Mathematical Statistics

2025

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

DECLARATION

I have read and understood the university's definition of plagiarism and cheating given in the Research Degrees Handbook. I declare that this thesis is my own work. I have acknowledged all results and quotations from other people's published or unpublished work.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only. Published by the University of Cape Town (UCT) in terms of the non-exclusive licence granted to UCT by the author.

Signature: _____ Signed by candidate _____ Date: _13/03/2025_

Ms Modiehi Violet Theletsane

Copyright © Modiehi Violet Theletsane 2025. All rights reserved.

ACKNOWLEDGEMENTS

I want to thank and praise Almighty God for the strength, resilience, patience, wisdom, and courage he instilled in me throughout undertaking this study. More importantly, I thank the Holy Spirit for leading and guiding me in times of hardship.

I sincerely thank my supervisor, Associate Professor Freedom Gumedze (University of Cape Town, South Africa), for his patience, professional guidance, and academic support throughout the writing of this thesis. He consistently allowed me to own this thesis but steered me in the right direction whenever he thought I needed it. Thank you for providing me with the IMPI dataset and valuable criticism.

Special gratitude is expressed to my mentor, Dr Isaac Singini (SAMRC), for his guidance and contribution in writing this thesis.

My beloved family and friends, thank you for your unfailing support and continuous encouragement during this journey.

Thank you all!

ABSTRACT

It is common in clinical or observational studies to record information repeatedly over time while observing a time-to-event outcome among subjects. Joint models for longitudinal and survival data simultaneously analyse repetitively measured outcomes and associated event times. They offer valuable applications in two contexts: accounting for time-varying covariates measured with error when concentrating on survival outcomes, and controlling for informative censoring when focusing on longitudinal outcomes. It has been nearly four decades since joint modelling was first developed. The main aim of this study was to investigate whether there is an association between multivariate longitudinal electrocardiogram (ECG) characteristics; i.e., ECG rate (ECGrate), ECG PR interval (ECGpri), ECG corrected QT interval (ECGqtc), and ECG QRS duration (ECGqrsd) on survival outcomes, death, constriction, and composite outcome (death, constriction, or cardiac tamponade, whichever occurs first) in the investigation of the management of pericarditis (IMPI) in a multi-centre clinical trial. The ECG characteristics from the IMPI trial were weighed on a continuous scale and were converted into categories to be clinically meaningful. Several approaches were taken towards joint modelling, with the first one being a two-stage joint model approach. The shared parameter joint model is another approach to joint modelling. This study considered univariate and multivariate shared parameter joint models of the longitudinal data and time-to-composite, time-to-death, and time-to-constriction event outcomes. Specifically, the study considered these models when the data were non-normal. The univariate analysis results suggested a weak association between the ECGrate and the risk of composite, death, and constriction event outcomes. However, there was a strong association between ECGpri and the risk of death and constriction, but there was no association in the composite event. Furthermore, there was no association between ECGqrs duration and the risk of either composite or death events; however, there was an association with constriction. Finally, there was no association between ECGqtc and the risk of either composite or death events; however, there was an association between the ECGqtc and constriction. The study utilised multivariate shared parameter joint model analysis to understand if there was an association between composite, death, and constriction survival outcomes. The model had four binary ECG longitudinal outcomes, which were modelled based on the binomial assumption using the generalised linear mixed-effects model. Parameter estimation was based on a Bayesian framework utilising the Markov Chain Monte Carlo technique, and convergency estimates were established. It was discovered that the association parameter for the ECGqtc, which determines how the longitudinal ECGqtc is related to the risk

of death, showed that there was an association. In contrast, the association parameter for ECGrate, ECGpri, and ECGqrsd was weak for the risk of composite, death, and constriction outcomes. The ECGqtc also revealed no association between the risk of composite and constriction event outcomes, respectively.

Keywords: Joint models, longitudinal, multivariate, survival

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µmol/l	Micromoles per litre
AI	Artificial intelligence
AIC	Akaike information criterion
AIDS	Acquired immunodeficiency syndrome
BIC	Bayesian information criterion
bpm	Beats per minute
cm	Centimetre(s)
COVID-19	Coronavirus disease of 2019
ECG	Electrocardiogram
ECGpri	Electrocardiogram PR interval
ECGqrsd	Electrocardiogram QRS duration
ECGqtc	Electrocardiogram corrected QT interval
ECGrate	Electrocardiogram rate
EM	Expectation-Maximisation
GHQ	Gauss-Hermite quadrature
GLMM	Generalised linear mixed-effects model
Hg	Millilitre(s) of mercury
HIV	Human immunodeficiency virus
HR	Hazard ratio
IMPI	Investigation of the Management of Pericarditis [Trial]
IQR	Interquartile range
kg	Kilogram(s)
LME	Linear mixed-effects
m.g.f.	Moment generating function
MCMC	Markov Chain Monte Carlo
mg	Milligram(s)
MIP	*Mycobacterium indicus pranii*
ml	Millilitre(s)
ML	Maximum likelihood
MLE	Maximum likelihood estimation
mm	Millimetre(s)
mm^3	Cubic millimetre(s)
ms	Millisecond(s)
MW	*Mycobacterium indicus pranii*

NYHA	New York Heart Association
PH	Proportional hazards
PM	Posterior mean
SD	Standard deviation
SE	Standard error
TB	Tuberculosis

CHAPTER 1:
INTRODUCTION

1.1 INTRODUCTION

This chapter introduces the study and outlines its objectives and goals. It also provides the motivation for the research and an overview of the thesis organisation to inform the intended audience of the current research. The background description of the research problem is provided in the next chapter.

Joint modelling, which has been an active research area for nearly four decades, characterises the relationship between longitudinal and time-to-event processes. It is common in clinical or observational studies to record information on the time-to-event outcome and the longitudinal outcome of the disease marker on subjects. These study designs primarily evaluate the longitudinal trend for the population as a whole or for a particular subject over time. As part of the data-collection process, other covariates may be included that are thought to influence the response variable.

Longitudinal and event processes were analysed separately in the past. For example, when interest lies in the time-to-event process, Cox regression with time-varying covariates or two-stage modelling has historically been employed. However, through the development of a statistical methodology, the two processes have been simultaneously analysed through a modelling framework, referred to as joint modelling, for longitudinal and time-to-event (survival) data (Henderson et al., 2000; Hsieh et al., 2006; Rizopoulos, 2010). Joint models for longitudinal and time-to-event data model the two processes simultaneously through the semi-parametric maximum likelihood (ML) approach and Bayesian method employing the Markov chain Monte Carlo (MCMC) technique (Chi & Ibrahim, 2006; R Brown & G Ibrahim, 2003; Xu & Zeger, 2001; Wang & Taylor, 2001). This new and active area of research can be utilised to improve inference about longitudinal profiles with informative dropout.

Joint models can be divided into two classes: latent class joint models and shared parameter joint models. The former model is useful in personalised medicine for predicting the occurrence of an event, while the latter assesses the strength of association between biomarkers and event risks (Henderson et al., 2000; Rizopoulos, 2012; Tsiatis & Davidian, 2001). This study primarily focused on the shared parameter model.

Additionally, the joint model has been extended to manage multivariate longitudinal data, competing risks, and recurrent events. This study examined the longitudinal multivariate data setting and competing risks. For countless applications, it may be of interest to jointly model longitudinal data with several random variables. This type of data is referred to as multivariate longitudinal data. In the statistical literature, it is argued that this extension is mathematically straightforward. However, in practice, it remains challenging to fit due to many random effects, which can also be computer intensive.

As part of time-to-event data analysis, an alternative event may occur apart from the event of interest, which is denoted to as competing risks. Competing risks describe events that either prelude the event of interest, prevent it from occurring, or change the likelihood of its occurrence. For example, in the Investigation of the Management of Pericarditis (IMPI) trial data, the competing risks are event death or constriction. The two universal models for analysing competing risk data are the sub-hazards model, also known as the Fine Gray model (Fine & Grey, 1999), and the cause-specific hazard model (Prentice, 1982). The Cox proportional hazards (PH) model is the base for both models. Sub-hazard models directly relate the covariate effect to the cumulative incident function, whereas cause-specific hazard models relate the covariate effect to the cause-specific incident function (Ha et al., 2016; Prentice, 1982).

In recent years, there has been a significant increase in longitudinal data modelling and competing risks in time-to-event data research. As a baseline model, linear mixed models were generally used for all longitudinal outcome sub-models. Several extensions have been proposed, including modelling the random effects heterogeneously, modelling multivariate longitudinal data, and non-parametric modelling of the random effects (Andrinopoulou et al., 2014; Andrinopoulou et al., 2017; Elashoff et al., 2010; Huang et al., 2011; Li et al., 2012). The association between sub-models has generally been captured through different functions of shared random effects (Li et al., 2012) or correlated random effects and frailty terms (Elashoff, 2008).

Most research in joint modelling has predominantly focused on continuous longitudinal responses. However, investigators often encounter various data types when conducting longitudinal studies, including continuous, binary, ordinal, and others. For example, Faucett et al. (1998) and Larsen (2004) developed joint models for binary longitudinal outcomes, while Pulkstenis et al. (Pulkstenis, 1998) and Albert (2000) have also applied them to missing data

cases. Many techniques in joint modelling literature are intended to cope with different types of longitudinal outcomes. The researcher therefore extended the linear mixed models for continuous data to non-continuous data using the generalised linear mixed-effects models (GLMMs) (Bogun et al., 2004; Engel & Keen, 1994; Laird & Ware, 1982; WHO, 2022). The focus of this study is on non-continuous data; specifically four binary longitudinal outcomes.

1.2 MOTIVATION

The multi-centre IMPI clinical trial, conducted approximately nine years ago, of which the main results were published in 2014, motivated the work in this study (Mayosi et al., 2014). The study's primary objective was to determine whether prednisolone and *Mycobacterium indicus pranii* (MIP) immunotherapy reduced the composite outcomes of death, constriction, or cardiac tamponade in patients with tuberculosis (TB) pericarditis with probable or definite outcomes.

The work presented in this thesis used the following variables from the IMPI trial data: electrocardiogram (ECG) rate (ECGrate), ECG PR interval (ECGpri), ECG QRS duration (ECGqrsd), and ECG corrected QT interval (ECGqtc) to clinically prove their importance in cardiology (Mayosi et al., 2014). Cardiologists use ECGs to diagnose and monitor cardiac conditions.

Hypertension, among many, is one of the leading causes of ECG abnormalities, which later harm cardiovascular health. Several studies in the literature have confirmed these findings (Bang & Okin, 2016; Ferro et al., 2021; Istolahti et al., 2021). A study by Lehtonen et al. (2016) specifically examined the prevalence of ECG abnormalities in hypertensive individuals and revealed that abnormalities increase with the severity of hypertension. A further finding of the study was that left ventricular hypertrophy remains the cornerstone of cardiovascular risk assessment in hypertensive patients. Moreover, it alerted physicians to look for the presence of prolonged QT interval and negative T-wave in hypertensive individuals, which is highly associated with the risk of heart disease. Adedinsewo et al. (2020) demonstrated the correctness of an artificial intelligence (AI)-enabled ECG to specifically identify patients presenting with acute dyspnoea who had left ventricular systolic dysfunction. The study revealed the accuracy of utilising the AI algorithm in the emergency room and a further high risk of mortality for patients with left ventricular systolic dysfunction with the added advantage of early linkage to cardiovascular care. Subsequently, McCullough et al. (2020) investigated the association

between ECG characteristics and death in patients with coronavirus disease (COVID-19). They achieved this by using univariate and multivariate logistic regression analysis. They found that patients with ECG discoveries of both left-sided heart disease and right-sided disease had a higher probability of death.

Despite ECGs being a powerful tool for diagnosing heart disease, incorrect interpretation can lead to inappropriate clinical decisions (Bogun et al., 2004; Masoudi et al., 2006; Viljoen et al., 2017;). It is for this reason that the above-mentioned ECG variables from the IMPI trial dataset were transformed from continuous to categorical variables to make the variables clinically meaningful, and they are now explained in terms of the normal versus abnormal rate of each ECG variable. The ECGrate represents the frequency of heart contraction beats per minute (bpm). It is clinically meaningful at the bedside, with a normal heart rate ranging between 60 to 100 bpm. Rates less than 60 and above 100 bpm are deemed abnormal.

An ECGpri is the time it takes for an electrical impulse to be generated by a pacemaker (sinoatrial node, i.e., red star in Figure 1.1) and then to be transmitted through both atria (i.e., atrial depolarisation, the blue waves in Figure 1.1) to the atrioventricular node (black oval dot in Figure 1.1). Figure 1.2 shows a visual representation of the process. The normal PR interval is 120 to 200 milliseconds (ms), and anything outside this interval is deemed abnormal. The abnormalities stem from when patients have a fast heart rate, which causes the ECGpri to be shortened, which can be prolonged to first-degree heart block (Viljoen et al., 2021).

The ECGqrsd complex represents ventricular depolarisation, which is the period the electrical impulse takes to pass through the left and right ventricles of the heart. It triggers the contraction of both ventricles' muscles. It is regarded as a central spike as visualised in Figure 1.3. The normal QRS width is 60 to 100 ms. However, when the QRS duration is longer, it is considered a wide and abnormal QRS complex, which the bundle branch block can likely cause.

Finally, the ECGqtc is the corrected QT interval that represents electrical activation and recovery of the ventricles; i.e., time it takes for an electrical impulse to travel through both ventricles, triggering ventricular contraction. After that, the ventricles return to their baseline electrical state. The QT interval is measured from the beginning of the QRS duration complex to the end of the T-wave; i.e., ventricular repolarisation, as shown in Figure 1.4 (Viljoen et al., 2021). If prolonged, many causes may be acquired, including disease states and multiple drugs.

Therefore, to be considered normal, the ECG QT interval must be less than 440 ms for men and less than 460 ms for women.

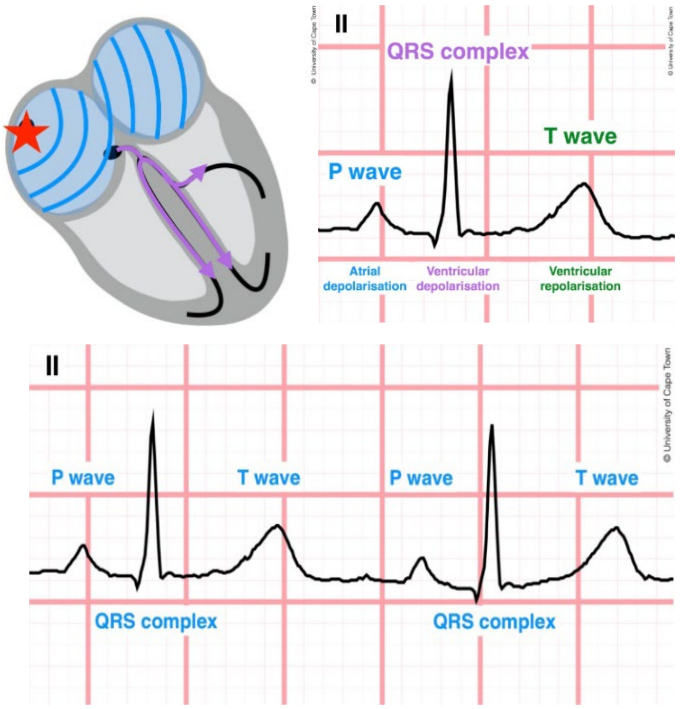

Figure 1.1: A visual representation of the ECGrate (Viljoen et al., 2021)

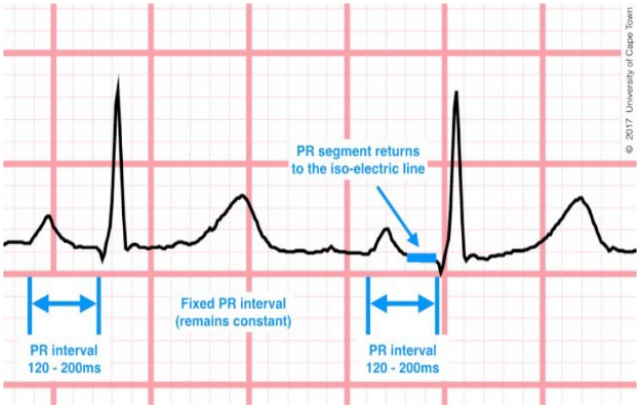

Figure 1.2: A visual representation of the ECGpri (Viljoen et al., 2021)

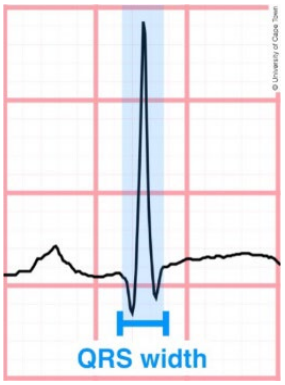

Figure 1.3: A visual representation of the ECGqrsd (Viljoen et al., 2021)

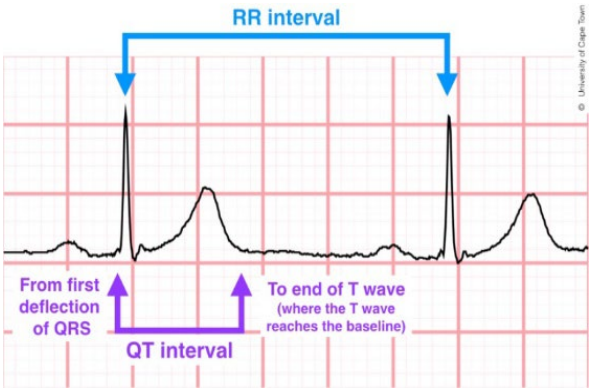

Figure 1.4: A visual representation of the ECGqtc (Viljoen et al., 2021)

In the current literature, the IMPI dataset analysis was only based on survival outcomes in joint modelling. Additionally, the published study did account for missingness, measurement error in the longitudinal outcome, and non-random dropout in the survival outcome. However, the concentration of the original study was on CD4 cell counts and not on the ECG characteristics variables.

1.3 STUDY AIM AND OBJECTIVES

This study aimed to assess whether the ECG variables, namely ECGrate, ECGpri, ECGqrsd, and ECGqtc, were individually or jointly associated with composite outcome, death, or constriction in TB pericarditis patients using the joint modelling framework.

The objectives of the study were to:

i. Investigate the association between the ECG characteristics and study outcomes using separate joint models, evaluate their performance, and illustrate them on the actual dataset from a multi-centre clinical trial (IMPI).
ii. Investigate the association between the ECG characteristics and study outcomes using separate joint models, taking into account competing risks.
iii. Investigate the association between the ECG characteristics and study outcomes using multivariate joint models.

1.4 CONTRIBUTION OF THE THESIS

The study is important because clinicians can use the identified ECG characteristics to be associated with the outcomes to identify patients at risk and therefore recommend appropriate medications or therapeutic interventions.

1.5 STRUCTURE OF THE THESIS

This thesis is organised as follows:

Chapter 1 presents the study's general introduction, motivation, and objectives. Chapter 2 provides the background and descriptive statistics of the IMPI clinical trial dataset. Chapter 3 reviews available literature on time-to-event (survival) analysis and the Cox PH model. Chapter 4 discusses relevant functions and standard methods in the time-to-event data analysis. Furthermore, the concept of joint modelling and its extensions are discussed, followed by the documentation used in the joint modelling framework. Chapter 5 introduces the multivariate joint model for longitudinal and survival data, formulation, and estimation (Rizopoulos, 2012). Chapter 6 presents the model fitting results and interpretation found after using each of the methodologies discussed in Chapter 4 and Chapter 5 on the IMPI real dataset. Finally, Chapter 7 presents the concluding facts discovered in this study, and makes suggestions for future research.

CHAPTER 2:
BACKGROUND TO THE RESEARCH PROBLEM AND INVESTIGATION OF THE MANAGEMENT OF PERICARDITIS (IMPI) TRIAL DATASET

2.1 INTRODUCTION

Chapter two provides a detailed description of the research problem. It summarises the key features of the IMPI clinical trial that was used to evaluate whether prednisolone and MIP reduced death, constriction, or cardiac tamponade among patients with TB pericarditis with possible or definite outcomes requiring pericardial drainage. In this study, the analysis was restricted to patients with ECG data. The chapter provides a brief exploratory data analysis. The next chapter discusses literature relevant to the study.

2.2 BACKGROUND TO THE RESEARCH PROBLEM

TB is one of the most severe health challenges globally. While its prevalence has declined, millions of people continue to be affected yearly, which makes it almost as common as the human immunodeficiency virus (HIV). The World Health Organization revealed in 2022 that around a quarter of the global population has been infected with TB, of which approximately 90% of people who develop TB each year are adults, with more cases occurring in men than in women (WHO, 2022). Since TB is a bacterial infection, the bacterium *Mycobacterium tuberculosis* is the primary cause of it. It usually affects the lungs (pulmonary TB) but can also affect other sites, such as the pericardial membrane. The pericardial membrane is the pericardium that covers the heart. It is therefore called TB pericarditis when it is infected with TB.

According to Mayosi et al. (2005), TB pericarditis accounts for approximately 1% of TB-related deaths and 1% to 2% of pulmonary TB cases. Due to this condition, there is a high incidence of pericardial effusion, cardiac tamponade, and constrictive pericarditis in emerging countries. Pericardial effusion is when fluid accumulates abnormally in the heart cavity, while cardiac tamponade occurs when blood builds up between the heart muscle and the pericardium, which causes pressure to build on the heart, and constrictive pericarditis is inflammation of the pericardium. A high risk of HIV infection is associated with TB pericarditis (Mayosi et al.,

2006). HIV is prevalent in almost half of the patients in sub-Saharan Africa with large pericardial effusions (Ntsekhe et al., 2003). A large pericarditis effusion is most common in patients with TB pericarditis in the Western Cape province (Reuter et al., 2005).

The treatment of TB pericarditis includes rifampicin, isoniazid, pyrazinamide, and ethambutol administered for six months, followed by pericardial drainage to relieve cardiac tamponade and pericardiectomy for pericardial constriction (Wiysonge et al., 2017). However, mortality caused by TB pericarditis remains high despite anti-TB chemotherapy (Mayosi et al., 2008). Furthermore, it was found that glucocorticoid therapy reduced the inflammatory response in TB pericarditis patients, improved outcomes, and decreased death risk due to reduced cardiac tamponade and pericardial constriction. However, it is unclear how effective adjunctive glucocorticoids are clinically (Mayosi et al., 2014). Nevertheless, Ntsekhe et al. (2003) found that adjunctive glucocorticoids were not guaranteed to reduce mortality or morbidity. It has been shown that glucocorticoids are effective for treating TB pericarditis, but that the number of events and patients included was very small, according to a meta-analysis involving randomised and controlled trials (Ntsekhe et al., 2008; Wiysonge et al., 2017). TB pericarditis patients could benefit from adjunctive prednisolone, and intradermal MIP could be effective in subduing inflammation and its sequelae in patients with TB pericarditis, according to Mayosi et al. (2014). The non-pathogenic, saprophytic, rapidly growing atypical MIP has demonstrated clinical benefits when applied as an intradermal heat-killed formulation to patients with leprosy, and it may also be beneficial in patients who are infected with HIV and pulmonary TB. It has been suggested that repeated doses of intradermal heart-killed MIP immunotherapy may lessen the inflammation response to TB and increase the CD4+ T-cell in HIV-positive individuals (Mathur, 2006).

2.3 DATA SOURCE

The IMPI clinical trial was a two-by-two factorial study conducted in nine African countries with multiple centres (Mayosi et al., 2014). The participants in the study were 18 years or older who had TB pericarditis. Patients were randomised to either an experimental or a control group during the trial, based on a random allocation process with stratification by the centre. The experimental group received the treatment, while the control group received a placebo. Firstly, they were assigned the prednisolone treatment for six weeks at a dose of 120 mg/day to 5 mg/day compared to the placebo. They were then randomly assigned to five injections of heat-

killed *Mycobacterium indicus pranii* (MIP or MW), at a dose of two 0.1 ml injections, and then subsequent 0.1 ml injections for comparison with the placebo. The clinical trial therefore consisted of Mw+P+, MW-P-, MW+P-, and MW-P+, where - indicates the placebo and + indicates the active treatment. An example would be P+MW+, which means the arm had active treatment for MW and prednisolone. The study included 1 400 patients, of whom approximately two-thirds were HIV positive, around three-quarters had heart disease of New York Heart Association (NYHA) class II or III, and approximately 90% had moderate to large pericardial effusions. The patients achieved a median follow-up of 600 days and high adherence and follow-up rates. After hospital discharge, data were collected at two weeks, four weeks, six weeks, three months, and six months, then every six months for two years, and every year after that up to four years.

Composite death, cardiac tamponade, or constrictive pericarditis was the primary efficacy outcomes of the study. It was concluded that there was no significant difference between the patients who received prednisolone and the placebo. Furthermore, the same dataset also measured the safety of immunomodulatory treatment. Among patients with HIV, prednisolone therapy was associated with an increased risk of many cancers, most of which were related to HIV (Mayosi et al., 2014).

The main aim of this study was to assess the effect of the ECG characteristics, namely ECGrate, ECGpri, ECGqtc, and ECGqrsd, which are the time-varying covariates, on survival outcomes, death, and pericardial constriction. The project examined the ECG individually and jointly versus death versus constriction and composite (death or constriction). Moreover, the project was considered suitable for using this data source because it had multivariate longitudinal and time event outcomes. The analysis was restricted to patients who only had ECG data. The study further focused on subjects with at least two ECG measurements, i.e., ECGrate, ECGpri, ECGqtc, or ECTqrsd (n = 1 197). The study only included subjects who were assigned prednisolone versus placebo treatment because the initial IMPI study flagged MIP as redundant. The trial collected the ECG measurements at baseline, two weeks, four weeks, six weeks, three months, six months, 12 months, 18 months, 24 months, and 48 months.

All the ethical protocols were observed, and approval was granted by the University of Cape Town Faculty of Health Science Human Research Ethics Committee, HREC (REF HREC 032/2009).

2.4 EXPLORATORY DATA ANALYSIS

The study enrolled 1 197 patients, of whom 845 were HIV positive, and 352 were HIV negative. Five hundred and ninety-three patients were randomised to a prednisolone arm, and 600 patients were randomised to a placebo arm. There were 255 patients among the 1 197 patients who experienced a composite event (i.e., death, cardiac tamponade, or constriction), 163 died, and 86 suffered cardiac arrest.

Thirty-three covariates were identified as clinically appropriate through a review of the covariates. A preliminary univariate analysis was conducted to further assess the effects of these 33 covariates on the response variables. It is worth noting that observations beyond the study period or lost to lack of follow-up were right-censored.

Missing values are common in medical datasets. As a result, this study assumed that data were missing at random and that censoring was non-informative.

2.4.1 Baseline characteristic distributions for the treatment group

The study used different statistical tests to compare baseline characteristic distributions. A Mann-Whitney test was used for continuous variables, while a chi-square or Fisher's exact test was used for categorical variables. The continuous covariates (see Table 2.1) and categorical covariates (see Tables 2.2 and 2.3) are summarised below. Hypothesis testing was done at a 5% significance level, whereby p-values of covariates ($p \geq 0.05$) indicated no significant differences in patients between the treatment groups at the baseline. Table 2.4 summarises the ECG baseline characteristics by treatment group, while Tables 2.5 to 2.7 show the summary characteristics for ECG baseline characteristics by study outcomes.

Table 2.1: Summary statistics for continuous baseline characteristics of patients randomised to either placebo or prednisolone group

Covariate	Placebo (N = 604) Median (IQR*)	Prednisolone (N = 593) Median (IQR)	Overall (N = 1 197) Median (IQR)	p-value
Age (years)	34.34 (15.82)	34.93 (17.80)	34.51 (16.44)	0.622
Weight (kg)	57.9 (15.0)	58.0 (15.7)	5.80 (1.5)	0.026
Duration of symptoms (days)	30 (22)	28 (21)	30 (21)	0.370

* IQR = interquartile range

Table 2.2: Summary statistics for categorical baseline characteristics of patients randomised to either placebo or prednisolone group

Covariate	Placebo (N = 604) N(%)	Prednisolone (N = 593) N(%)	Overall (N = 1 197) N(%)	p-value
Gender (sex)				1.000
Female	267 (22.31)	263 (21.97)	530 (44.28)	
Male	337 (28.15)	330 (27.57)	667 (55.72)	
NYHA class				0.580
I	122 (10.19)	129 (10.78)	251 (20.97)	
II	277 (23.14)	277 (23.14)	554 (46.28)	
III	151 (12.61)	129 (10.78)	280 (23.39)	
IV	54 (4.51)	58 (4.85)	112 (9.36)	
Country				0.807
Other	138 (11.53)	131 (10.94)	269 (22.47)	
South Africa	466 (38.93)	462 (38.60)	928 (77.53)	
Pericardiocentesis at randomisation				0.656
No	248 (20.72)	235 (19.63)	483 (40.35)	
Yes	356 (29.74)	358 (29.91)	714 (59.65)	
Antiretroviral status at study entry				0.048
Unknown	180 (15.04)	193 (16.12)	373 (31.16)	
No	317 (26.48)	325 (27.15)	642 (53.63)	
Yes	107 (8.94)	75 (6.27)	182 (15.20)	
Palpable pulsus paradoxus				0.049
No	483 (40.35)	445 (37.18)	928 (77.53)	
Yes	121 (10.11)	148 (12.36)	269 (22.47)	
Peripheral oedema				0.253
No	375 (29.14)	388 (29.79)	763 (58.93)	
Yes	229 (20.43)	205 (20.64)	434 (41.07)	
Systolic blood pressure ≤90 mm ml of mercury (Hg)	29 (2.42)	42 (3.51)	71 (5.94)	0.120
>90 mm Hg	575 (48.08)	550 (45.99)	1 125 (94.06)	
Heart rate ≤100	255 (21.34)	251 (21.00)	506 (42.34)	0.976
>100	349 (29.21)	340 (28.45)	689 (57.66)	
Haemoglobin ≤10	317 (26.62)	330 (27.71)	647 (54.32)	0.392
>10	281 (23.59)	263 (22.08)	544 (45.68)	
White blood count				0.051
≤10 mm^3	566 (45.49)	540 (47.42)	1 106 (92.91)	
>10 mm^3	35 (4.08)	53 (3.01)	88 (7.09)	

Table 2.3: Summary statistics for categorical baseline characteristics continued

Covariate	Placebo (N = 604) N(%)	Prednisolone (N = 593) N(%)	Overall (N = 1 197) N(%)	p-value
Creatinine≤105 umol/l	508 (45.72)	489 (44.01)	997 (89.74)	0.327
>105 umol/l	52 (4.68)	62 (5.58)	114 (10.26)	
Left ventricular ejection fraction ≤55%	1 (0.08)	1 (0.08)	2 (0.17)	1.000
>55%	603 (50.38)	592 (49.46)	1 195 (99.83)	
Effusion size	383 (32.85)	363 (31.13)	746 (63.98)	0.445
Large ≥2 cm				
Medium 1-2 cm	148 (12.69)	160 (13.72)	308 (26.42)	
Small <1 cm	61 (5.23)	51 (4.37)	112 (9.61)	
Tamponade at study entry				1.000
No	216 (23.53)	215 (23.42)	431 (46.95)	
Yes	244 (26.58)	243 (26.47)	487 (53.05)	
Constriction at study entry				0.026
No	262 (28.32)	231 (24.97)	493 (53.30)	
Yes	197 (21.30)	235 (25.41)	432 (46.70)	
HIV status				0.303
HIV-	169 (14.12)	183 (15.29)	352 (29.41)	
HIV+	435 (36.34)	410 (34.25)	845 (70.59)	
Chest X-ray pulmonary infiltrate				0.556
No	379 (34.36)	347 (31.46)	726 (65.82)	
Yes	189 (17.14)	188 (17.04)	377 (34.18)	
Atrial fibrillation on ECG				1.000
No	489 (50.31)	480 (49.38)	969 (99.69)	
Yes	2 (0.21)	1 (0.10)	3 (0.31)	
Definite TB pericarditis status				0.950
No	506 (42.27)	495 (41.35)	1 001 (83.63)	
Yes	98 (8.19)	98 (8.19)	196 (16.37)	
Pericardial thickness				
Normal (≤2 mm)	603	592	1 195	

Table 2.4: Summary statistics for electrocardiogram (ECG) baseline characteristics

Covariate	Placebo (N = 604) N(%)	Prednisolone (N = 593) N(%)	Overall (N = 1 197) N(%)	p-value
ECGrate				0.425
Abnormal	248 (20.72)	258 (21.55)	506 (42.27)	
Normal	356 (29.74)	335 (27.99)	691 (57.73)	
ECGpri				0.910
Abnormal	585 (48.87)	576 (48.12)	1 161 (96.99)	
Normal	19 (1.59)	17 (1.42)	36 (3.01)	
ECGqrsd				0.069
Abnormal	593 (49.54)	571 (47.70)	1 164 (97.24)	
Normal	11 (0.92)	22 (1.84)	33 (2.76)	
ECGqtc				0.491
Abnormal	511 (42.69)	492 (41.10)	1 003 (83.79)	
Normal	93 (7.77)	101 (8.44)	194 (16.21)	

Table 2.5: Summary statistics for ECG baseline characteristics by composite outcome

Covariate	No composite (N = 942) N(%)	Composite (N = 255) N(%)	Overall (N = 1 197) N(%)	p-value
ECGrate				0.006
Abnormal	418 (34.92)	88 (7.35)	506 (42.27)	
Normal	524 (43.78)	167 (13.95)	691 (57.73)	
ECGpri				0.242
Abnormal	917 (76.61)	244 (20.38)	1 161 (96.99)	
Normal	25 (2.09)	11 (0.92)	36 (3.01)	
ECGqrsd				0.287
Abnormal	919 (76.78)	245 (20.47)	1 164 (97.24)	
Normal	23 (1.92)	10 (0.84)	33 (2.76)	
ECGqtc				0.001
Abnormal	807 (67.42)	196 (16.37)	1 003 (83.79)	
Normal	135 (11.28)	59 (4.93)	194 (16.21)	

Table 2.6: Summary statistics for ECG baseline characteristics by death outcome

Covariate	No death (N = 1 034) N(%)	Death (N = 163) N(%)	Overall (N = 1 197) N(%)	p-value
ECGrate				0.022
Abnormal	451 (37.68)	55 (4.59)	506 (42.27)	
Normal	583 (48.71)	108 (9.02)	691 (57.73)	
ECGpri				0.431
Abnormal	1 005 (83.96)	156 (13.03)	1 161 (96.99)	
Normal	29 (2.42)	7 (0.58)	36 (3.01)	
ECGqrsd				0.039
Abnormal	1 010 (84.38)	154 (12.87)	1 164 (97.24)	
Normal	24 (2.01)	9 (0.75)	33 (2.76)	
ECGqtc				0.166
Abnormal	875 (73.10)	128 (10.69)	1 003 (83.79)	
Normal	159 (13.28)	35 (2.92)	194 (16.21)	

Table 2.7: Summary statistics for ECG baseline characteristics by constriction outcome

Covariate	No constriction (N = 1 111) N(%)	Constriction (N = 86) N(%)	Overall (N = 1 197) N(%)	p-value
ECGrate				0.045
Abnormal	479 (40.02)	27 (2.26)	506 (42.27)	
Normal	632 (52.80)	59 (4.93)	691 (57.73)	
ECGpri				0.477
Abnormal	1 076 (89.89)	85 (7.10)	1 161 (96.99)	
Normal	35 (2.92)	1 (0.08)	36 (3.01)	
ECGqrsd				0.929
Abnormal	1 081 (90.31)	83 (6.93)	1 164 (97.24)	
Normal	30 (2.51)	3 (0.25)	33 (2.76)	
ECGqtc				0.491
Abnormal	936 (78.20)	67 (5.60)	1 003 (83.79)	
Normal	175 (14.62)	19 (1.59)	194 (16.21)	

2.4.2 ECG characteristics over time

Furthermore, this analysis considered the ECG characteristics (i.e., ECGrate, ECGpri, ECGqtc, and ECGqrsd) as the primary longitudinal outcomes. Table 2.4 presented the baseline categorical ECG characteristics. The ECG measurements reduced steadily with time from baseline until 11 months, as displayed in Figures 2.1 to 2.4. It is observed that one would model ECGrate displayed in Figure 2.1 using a linear model but not ECGpri, for example. The GLMM was thus used to model all the ECG characteristics. The treatment group analysis was restricted to only the prednisolone and placebo groups. The reason for this is because the comparison

between MIP and the placebo was stopped early in February 2013 due to futility (Mayosi et al., 2014).

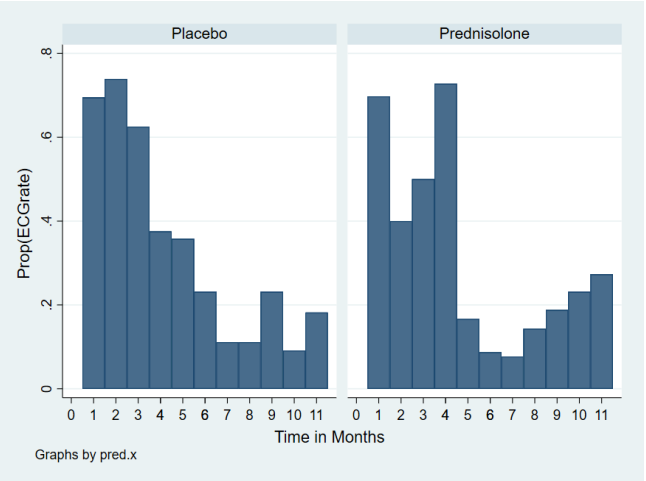

Figure 2.1: ECGrate proportion by month visit

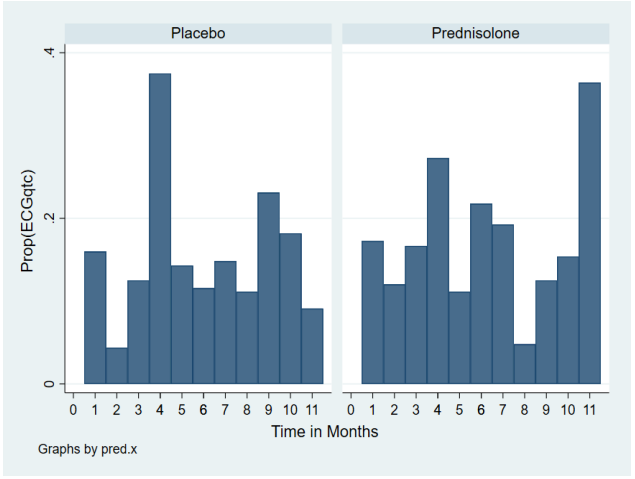

Figure 2.2: ECGqtc proportion by month visit

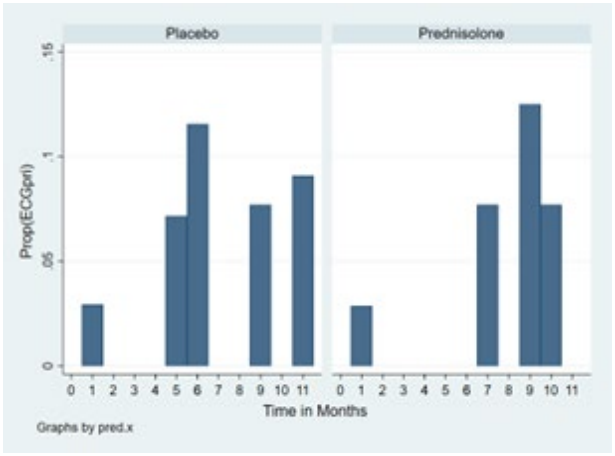

Figure 2.3: ECGpri proportion by month visit

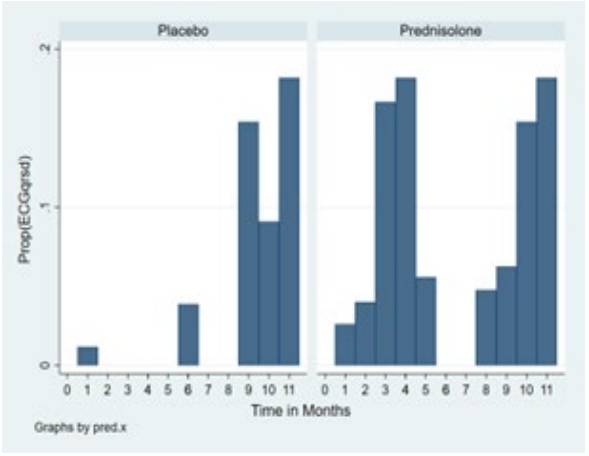

Figure 2.4: ECGqrsd proportion by month visit

CHAPTER 3:
LITERATURE REVIEW

3.1 INTRODUCTION

This chapter discusses literature on essential concepts of survival analysis, the Cox PH model and some extensions of it, as well as joint modelling for longitudinal and time-to-event studies using the shared parameter model and the latent class model. The chapter further discusses multivariate joint modelling for longitudinal and time-to-event studies when interest is in the binary longitudinal outcome. The theory of joint models for longitudinal and the time-to-event data employed in this study are discussed in the next chapter.

3.2 ESSENTIAL FUNCTIONS IN TIME-TO-EVENT ANALYSIS

Survival analysis is a statistical procedure for describing the time to the occurrence of an event. It is common in medical studies that patients are followed up over a specific period during an investigation of a particular drug. The research of such studies observes an event of interest that could be either death, the onset of a disease, or dropping out during the study. Censoring is the defining characteristic of survival data. Firstly, it can occur if there is a loss to follow-up or non-occurrence of the outcome event during the observation period or at the end of the trial, known as right censoring, or if the patient was at risk for disease before entering the study, known as left censoring. Lastly is interval censoring, where the monitoring assessment is done at a periodic interval, and the time to event is only known for a specified period (Prinja & Gupta, 2010). This study focused mainly on survival analysis methods applied to biomedical data.

The time-to-event data were analysed using the following functions: survival function, hazard function, probability density function, and cumulative hazard function. The non-negative random variable associated with the individual survival time, denoted by T, is the probability that an individual random selection will have a survival time less than or equal to some value, t. Formally, it is given as:

$$F(t) = P(T < t)$$
$$= \int_0^t f(v)dv,$$

where $f(v)$ is the probability density function of some variable v. The survival function represents an individual's survival distribution from the beginning of the trial at time (t) and beyond without experiencing the event of interest. Formally, it is defined as the probability of survival time greater than a specific value (t) given by:

$$S(t) = P(T \geq t)$$
$$= 1 - F(t).$$

The hazard function, $h(t)$, known as the instantaneous failure rate or the force of mortality, gives the risk of event at a particular time t. It is the conditional probability that an individual has survived a particular time t while observing an event of interest:

$$h(t) = \lim_{\Delta t \to 0} \left\{ \frac{Pr(t < T < t + \Delta t \mid T > t)}{\Delta t} \right\},$$

where Δt is the instantaneous change in time t. The Nelson-Aalen estimator determines the cumulative hazard function by adding up the quotients of observed events and the number of subjects at risk; that is, by summarising all the observed events and the number of individuals at risk (Nelson, 1969):

$$\widetilde{\Lambda}\,(t) = \sum_{j=1}^{k} \frac{d_j}{n_j}.$$

A direct relationship exists between the probability density function $f(t)$, the survival function $S(t)$, and the hazard function $h(t)$ as follows:

$$h(t) = \frac{f(t)}{s(t)}.$$

Similarly,

$$S(t) = 1 - F(t) = \exp\big(-\Lambda(t)\big).$$

3.3 ANALYSIS OF TIME-TO-EVENT DATA

The evolution of covariates in the time-to-event data can be analysed using parametric, semi-parametric, and non-parametric methods. Analysing time-to-event data aims to model the survival time and describe the survival experience. Life tables and Kaplan and Meier are the two descriptive time-to-event data-analysis methods. A life table is a simple method

traditionally used by actuaries. In comparison, Kaplan and Meier is the most common method used by statisticians, which estimates the distribution of survival (Kaplan & Meier, 1958). This method focuses on the survival time curve of all subjects before the event of interest without considering the clinicians' attention within a fixed-cycle interval (Rich et al., 2010). Its main drawback is its inability to estimate the survival distribution while adjusting for covariates. In current data, we often wish to compare the survival experience of two or more patients. This can be achieved by obtaining the survival proportion of each group. However, the challenge will only provide the estimate at an arbitrary time point. Mentel (1966) introduced the log-rank test that intends to assist in comparing the survival distribution of two or more individuals and can therefore be regarded as a modified chi-square test (Lai et al., 2013). In theory, it is a hypothesis test only used when the data have been censored appropriately.

3.4 COX PROPORTIONAL HAZARDS (PH) MODEL

The Cox PH model is a regression model commonly known for modelling time-to-event data. Cox (1972) introduced the model, which assists in finding the association between demographic variables such as sex, age, etc., and explanatory variables (Cox, 1972). It is a semi-parametric model that assumes that covariates' effects are constant over time; the variable of interest, often censored, is non-negative (Moore, 2016). The model was rendered as follows by Cox (1972):

$$h_i(t) = h_0(t)\exp(\beta' x_i),$$

where $h_0(t)$ is the non-negative function of time known as baseline hazard function such that the vectors of explanatory variables are $x = 0$, $\beta' x_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}$, $\beta' = (\beta_1 \ldots \beta_p)$ that are unknown regression coefficients of the p covariates $X_1 \ldots X_p$. x_i is the vector of covariates for the i^{th} individual, $i = 1 \ldots n$. The exponential coefficients e^{β_j} are called the hazard ratios (HRs). The equation can be re-expressed as a linear model of the log ratio of the hazards as:

$$\frac{h(t\,|x_i)}{h(t\,|x_j)} = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p.$$

3.4.1 Parameter estimation in the Cox PH model

Parameter estimation of the unknown β coefficients in the Cox PH model can be found using the ML method. The method is performed by specifying the likelihood function of the sample data to be maximised; hence estimation using partial likelihood (Lemeshow et al., 2011). Partial likelihood only depends on the order of the deaths. Because of this, inferences about the explanatory variables are based on survival times only (Collet, 2003). There are detailed discussions of ML methods in survival analysis in the classic literature, such as by Kalbfleisch and Prentice (Kalbfleisch & Prentice, 2002).

Cox (1972) formed the bases of the partial likelihood function and further discussed the modification to suit nuisance parameters that may arise in survival data by using the dimension reduction technique as follows (Cox, 1975):

$$L(\beta) = \prod_{j=1}^{r} \left\{ \frac{\exp(\beta' x_{(j)})}{\sum_{I \varepsilon R_{(t_i)}} \exp(\beta' xl)} \right\}^{\delta_i},$$

where the observed data are right-censored such that $t_1 < t_2 < \cdots < t_{(r)}$ are ordered r death times with a group of individuals at risk denoted by $R(t_{(j)})$ before time $t_{(j)}$, so that $x_{(j)}$ gives a vector of covariates describing the individuals who die within the specified interval, and lastly δ_i, which denotes the event indicator for r observed survival times. The unknown β-parameters, which are estimated using the maximum likelihood estimation (MLE) method in the PH model, can be found by maximising the log-likelihood function. Consequently, the log-likelihood, which is computationally simpler, corresponds to:

$$\log L(\beta) = \sum_{i=1}^{n} \delta_i \left\{ \beta' x_i - log \sum_{I \varepsilon R(t_i)} \exp(\beta' x_i) \right\}.$$

Statistical packages such as R, STATA, and SAS have packages available to perform maximisation using the Newton-Raphson procedure.

3.4.2 Diagnostics for the Cox PH model

Statistical models mainly rely on assumptions. As a result, diagnostic techniques can be employed to establish whether the fitted model is adequate. In addition, the inferential conclusions based on a fitted model will be as robust and valid as possible if these procedures

are performed thoroughly and diligently (Lemeshow et al., 2011). According to the standard Cox PH regression model, PHs are a fundamental assumption (Cox, 1972; Collet, 2003; Hiller et al., 2015; Lemeshow et al., 2011). It is thus vital to investigate specific aspects of the model fit to ascertain the PH assumption, such as choosing explanatory variables and their functional forms and identifying outliers.

Residuals are a common diagnostic technique to assess the model's goodness of fit (Collet, 2003; Xue et al., 2017). In Cox modelling diagnostics, Schoenfeld residuals are used to determine if the proportionality assumption is violated. Martingale residuals are used to determine the functional form of a covariate. In addition, the Martingale plots are another way to assess the functional form of the covariates. However, there are limitations to this method. Therneau et al. (1990) proposed using the deviance residual to decide which observations are outliers. Other diagnostics include examining the overall fit of the survival model using the Cox-Snell residuals. The Cox-Snell residuals must be exponentially distributed to ensure a good model fit. In this study, the Cox-Snell residuals was used to assess various survival models.

As an alternative to the standard techniques mentioned above, graphical methods exist to assess the PH assumption for categorical or continuous variables with a few levels. However, it becomes challenging when the variables have many levels (Arjas, 1988; Cox, 1979; Xue et al., 2017). Hence, Therneau and Grambsch (2000) suggested plotting cumulative Schoenfeld residuals versus survival times. Nonetheless, the plots are challenging to read.

3.5 THE EXTENDED COX PH MODEL

Thus far, the standard Cox PH model that has been discussed only incorporates values of the explanatory variables recorded at baseline, assuming that the covariates were constant over time. Moreover, the PH assumption implies that the HR is constant over time; it is not time-dependent. However, most survival data arise from repeated longitudinal measurements and multiple explanatory variables observed over some time, and these variables vary throughout the study. Hence, a violation of the PH may occur in some instances with such variables. These variables are called explanatory variables, such as time-varying covariates or time-dependent covariates. Therefore, the extended Cox PH model to handle time-dependent covariates is now considered. This model, also known as the Anderson and Gill model, is most effective when the evolution of covariates does not have a proportional effect on the hazard function by using

the counting process formulation (Andersen, 1982; Chiou et al., 2014). The model formulation is presented in Chapter 4.

The stratified Cox model is another method for dealing with the violation of the PH assumption for a particular covariate. In this model, these covariates are incorporated as stratification factors, while those that meet the assumption are included as regressors. Mieno et al. (2011) ascertain that this technique is valuable in the presence of categorical predictors, which cause non-proportionality. There are also models that deal with the accelerated failure time, frailty models, and advanced models to deal with competing risks and multistate models (Sarkar et al., 2017). The next section discusses the advanced models used in this study.

3.6 JOINT MODELLING FOR LONGITUDINAL AND TIME-TO-EVENT DATA

The extended Cox PH model discussed in Section 3.5 only handles exogenous time-varying covariates but fails to account for the longitudinal relationship. Hence, the joint modelling framework for longitudinal and time-to-event data is an alternative approach, especially when the main interest is the association between survival experience and the endogenous time-varying covariates (Faucett & Thomas, 1996; Henderson, 2000; Tsiatis & Davidian, 2004; Wulfsohn & Tsiatis, 1997). Joint modelling has become the most popular technique to model the longitudinal measurement and time-to-event outcome simultaneously and has gained much attention from researchers in recent years. The advantage behind the method is the ability to minimise bias estimates on influential observations and to increase treatment efficacy and other prognostic factors. Several authors in the current literature cover comprehensive reviews focusing on jointly modelling the CD4 counts and the time an individual contracts acquired immunodeficiency syndrome (AIDS) (Hogan et al., 1998; Tsiatis & Davidian, 2004; Tsiatis et al., 1995; Self & Pawitan, 1992). There are various approaches to modelling the relationship between longitudinal development characteristics and time-to-event outcomes. A simple way is to fit an extended Cox model with longitudinal measures as time-varying covariates. Tsiatis et al. (1995) described this modelling approach as associated with problems such as underestimating the true association between the hazard and underlying response because of measurement errors (Prentice, 1982). The presence of measurement errors yields biased estimates. Another recent approach is two-stage joint modelling. In the first stage, the time-varying covariates are fitted in the linear mixed-effects (LME) model and the fitted values in the model are extracted. In the second stage, an extended Cox model with longitudinal measurements taken at different visits replaced by fitted values is included (Boycott & Taylor,

1998; Dafini & Tsiatis, 1998; Self & Pawitan, 1992; Tsiatis et al., 1995). Little attention has been paid to this approach in the literature. Perhaps because it still does not account for non-random dropout on the influential observations.

Consequently, Sweeting and Thompson (Sweeting & Thompson, 2012) argue that there is uncertainty in the estimates obtained in the first stage and therefore conclude that they are biased. Mauff et al. (2020) recently conducted a corrected two-stage approach that substantially reduces bias in the estimates and daunting computational time. The authors suggest a correction factor based on the importance of sampling theory (Press et al., 2007). However, the proposed approach has not been attempted much in the literature. The problem of unbiased estimates therefore persists. As a result, a new methodology for the joint distribution of the longitudinal and survival outcomes emerged (Rizopoulos, 2010; Rizopoulos, 2012). This unique and active approach in the statistics community, known as joint modelling, employs a joint maximisation of the likelihood using both the covariate process and the survival data. In addition, estimating the covariate process and risk of failure has proven to account for non-random dropout and provides valid inferences. There are currently two core joint modelling approaches in medical research: the shared parameter model, which is the focus of this study, and the latent class model. The shared parameter model uses the mixed-effects model for both the longitudinal process and time to dropout (Follmann & Wu, 1995). It is further assumed that both the longitudinal outcomes and survival use random effects to consider the association between these two outcomes. The papers of Tsiatis and Davidian (2004) and Yu et al. (2004) provide excellent overviews of this buzzing area of biostatistics and statistics research. Chapter 4 elaborates on the model formulation.

In contrast to the shared model, the joint latent class model captures the association between the risk hazard and the biomarker evolution by latent class structure (Lin et al., 2002; Proust-Lima & Taylor, 2009; Sartor et al., 1997). The authors argue that the use of class models is effective in studying prostate cancer recurrence. Specifically, they developed a tool to assist them in early detection of the resurface of prostate cancer in a cohort study of prostate cancer patients treated with radiotherapy. So far, only the standard joint models have been discussed, which only determine the association between survival outcome and a single longitudinal biomarker. However, several biomarkers may be relevant to the event of interest, especially in medical settings. Extending the univariate joint model case by incorporating multiple longitudinal biomarkers will therefore enable a better understanding of the underlying disease dynamics. In addition to the multivariate case, other extensions such as the ability to include

longitudinal outcomes of different types, competing risks, multiple continuous outcomes, recurrent and terminal events, and prediction, among others, are combined (Barrett & Su, 2017; Krol et al., 2016; Li et al., 2009; Lin et al., 2002; Rizopoulos et al., 2008).

This study uses Bayesian methods for fitting joint models. Bayesian approaches have also gained much attention in the literature in simultaneously addressing the longitudinal and time-to-event outcomes. Lawrence et al. (2015) reviewed this approach, focusing on the frequentist paradigm. In particular, the authors looked at case studies when multivariate longitudinal observations are taken into account, and multiple events are accounted for by integrating these random effects. Alvares and Rubio (2021) introduced a numerically tractable formulation of Bayesian joint models. Their method linked the longitudinal and survival processes by sharing the fixed and random effects. The reason was to include the nonlinear and time-dependent influences using the RStan package in R. In addition, the authors adopted MCMC for model parameter estimation to facilitate direct sampling from the posterior distributions of parameters. A parametric baseline hazard function was assumed for the survival component to avoid numerical integration and support multiple hazard models.

Rue et al. (2017) proposed Bayesian joint modelling for bivariate longitudinal and competing risk data. Using a shared parameter model, they model the two longitudinal and competing risk survival processes in the intensive care unit setting. In the longitudinal process, a mixed effects model is used for the sequential organ failure assessment score, while a mixed effects beta regression model is used for the index of asynchronies score. A survival process defines the competing risks of patients dying or being discharged alive through a cause-specific hazard model. As a result of this approach, the longitudinal biomarkers are linked to the time-to-event outcomes through shared latent random effects. The authors utilise the R-INLA package, which provides a fast and reliable inference technique for estimating parameters in the Bayesian paradigm when applying joint models to complex multivariate data. The work of Brilleman et al. (2019) addresses the application of a joint modelling approach to multilevel hierarchical data. They examined the relationship between tumour burden and disease progression in cancer patients. The authors used a hierarchical joint model with three levels in their approach. In level one, they accounted for repeated measurements within lesions (level two), which are then nested within patients (level three). They also employ Bayesian estimation, which facilitates the analysis of fixed effects (population-level parameters) and random effects (individual deviations).

Furthermore, in this study, the event of interest from the IMPI trial dataset was the composite outcome of death, constriction, or cardiac tamponade. In the case of constriction being the primary event of interest, other competing events such as death and cardiac tamponade may end up censoring it. Similarly, if cardiac tamponade is selected as the primary event of interest, constriction or death may compete with it. Therefore, when competing risks are present, caution should be exercised when estimating the probability of interest of events.

CHAPTER 4:
STANDARD JOINT MODELS FOR LONGITUDINAL AND TIME-TO-EVENT DATA

4.1 INTRODUCTION

This chapter discusses the theory of joint models for longitudinal and time-to-event data employed in this study. It is followed by the theory of joint models for longitudinal and time-to-event data when the data are non-normal. The next chapter extends the standard joint models for longitudinal and time-to-event data to a multivariate setting.

Modelling data with time-dependent covariates involves a variety of methods, especially when the aim is to jointly model longitudinal and time-to-event data. One can either use the extended Cox model, the two-stage joint model, or even evaluate the joint model of shared random effects using the shared parameter joint modelling of two processes. This discussion summarises the evolution of these joint models for longitudinal and survival data.

This chapter is two-fold. It starts by introducing the LME model as a building block to understand the joint modelling framework and discusses the methods mentioned above in detail and their extensions. Lastly, the joint models for the multiple longitudinal responses are discussed.

4.2 LINEAR MIXED-EFFECTS (LME) MODEL

This section introduces the LME model, which is the sub-model in the joint modelling framework by incorporating longitudinal data and event time data. It is common in medical studies to have data that arise from subjects (e.g., human beings or animals) that are followed up over time. Such data are known as longitudinal data, and their main aim is to investigate cross-sectional and longitudinal effects. The cross-sectional effect shows if a treatment differs on average at specific time points, while the longitudinal effect indicates the presence of differences in treatment effects (Rizopoulos, 2012). Longitudinal data analysis is based on the idea that each individual within a population has a subject-specific mean response profile; i.e., a unique functional form over time. The IMPI trial dataset has the longitudinal measurements of the ECG done at baseline and followed up for a minimum of six months for each subject (Chishala, 2016). Standard statistical methods such as linear regression and t-tests based on the

independence assumption and ordinary least squares estimation are therefore not adequate for analysing longitudinal data since subjects are positively correlated. It is for this reason that correlation must be taken into consideration to draw valid conclusions. Several techniques in the literature can take this into account, including generalised estimation equations. However, LME models are best suited to analyse longitudinal data since they account for both within- and between-subject correlations through random effects. In the joint modelling for longitudinal and time-to-event data framework (Rizopoulos, 2012), the mixed-effects models are used to predict subject-specific response profiles and estimate parameters that describe the mean response changes in the population. Moreover, the mixed-effects models assume that the longitudinal responses are independent and conditional on random effects. Additionally, these models are capable of handling data that are not balanced, and random effects are used to parsimoniously account for the correlation between repeated measurements.

4.2.1 Formulation of the LME model

The LME model has the following form (Harville, 1977; Laird & Ware, 1982; Thijs et al., 2000):

$$\boldsymbol{y_i} = \boldsymbol{X}_i\beta + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_{i,}\ i = 1, \dots, M$$
$$\boldsymbol{b}_i \sim N(\boldsymbol{0}, \sum), \qquad \boldsymbol{\varepsilon}_i \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$$

where $\boldsymbol{y_i}$ denotes the n_i-dimensional response vector for the ith group, $\boldsymbol{X_i}$ is $n_i \times p$ design matrix of fixed effects, while $\boldsymbol{Z_i}$ is $n_i \times q$ design matrix of random effects. β is the p-dimensional vector of fixed effects, $\boldsymbol{b_i}$ is the q-dimensional vector of random effects such that $q = \sum_{i=1}^{b} q_i$, with the assumption that $E(b_i) = 0$, ϵ_i is the n_i-dimensional vector of within-group errors assuming that $E(\epsilon_i) = 0$. Furthermore, b_i and ϵ_i both have independent and multivariate normal distribution in the following way:

$$\begin{bmatrix} b \\ \epsilon \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} G(\gamma) & 0 \\ 0 & \boldsymbol{R}(\boldsymbol{\rho}) \end{bmatrix}\right),$$

where γ and ρ are respectively r × 1 and s × 1 (with $s \leq \frac{n(n+1)}{2}$) vectors of unknown variance parameters corresponding to b and ϵ. If the random terms are correlated, then the dimensions of γ may exceed q i.e. γ may be of dimensions $r \leq \frac{q(q+1)}{2}$.

The variance-covariance matrix of the response vector, $\boldsymbol{y_i}$, can be expressed as follows (Patterson & Thompson, 1971):

$$Var(\boldsymbol{y_i}) = \sigma^2(\boldsymbol{Z_i G Z_i^T} + \boldsymbol{R}),$$

where $\boldsymbol{Z_i G Z_i^T}$ is the component of the random effect and $\boldsymbol{R}$ is the within-group component whereby both the components model correlation and heteroscedasticity. The variance-covariance matrix of the response variables can sometimes be modelled directly using the within-group component R to adjust for correlation among observations without incorporating random effects. A logarithmic likelihood maximisation is the most widely used estimation technique for linear models.

4.2.2 Parameter estimation for the LME model

This subsection concentrates on the general methods used to estimate the parameters in the LME model. The focus is on the ML and restricted ML. In the literature, several authors have investigated various estimation methods (Searle, 1995; Vonesh & Chinchilli, 1997). Simple linear models have primarily used least squares for parameter estimation, which assumes that the value with the least variance is the best one. An extension of this principle can be made to linear mixed modelling.

4.3 GENERALISED LINEAR MIXED-EFFECTS MODELS (GLMMs)

This section discusses the case when data are not normally distributed.

4.3.1 Model formulation

GLMMs are an extension of LME models (as discussed in Section 4.2) as they allow the response variable, Y_{ij}, to follow any distribution from the exponential family of distributions (McCulloch, 2003; McCulloch, 2001). LME models are frequently used when observations are continuous. However, in actual practice, the observations are either binary or categorical. For example, in the IMPI trial dataset, the ECG characteristics, which were the primary covariates of the project, take on the values of 0 or 1, whereby 0 means normal and 1 means abnormal. Therefore, they are categorical random variables. The most distinguishing factor of the GLMM is the relaxation of the independence assumption between observations that LME models are based on. This is achieved by adding the random effects to the fixed effects, which further

allows the models to capture the dependence structure that some grouping variables impose. As a result, GLMMs allow regression parameters to vary with any grouping variables suspected of causing responses to differ. Unlike LME models, GLMMs comprise various models that include normal, binomial, Poisson, and multinomial as special cases. These models therefore apply to cases where the observations may not be continuous. The GLMM general model definition is given as follows: The conditional probability density function of members of the exponential family is dependent on $\mathbf{b}_i$ and is given as:

$$\begin{aligned} f\left(y_{ij}\middle|\boldsymbol{b_i}\right) &= exp\left[\emptyset^{-1}\left\{y_{ij}\theta_{ij} - c(\theta_{ij})\right\} + d(y_{ij,}\emptyset)\right] \\ E\left(y_{ij}|\boldsymbol{b_i}\right) &= \mu_{ij} = g^{-1}\left\{\boldsymbol{X'_{ij}\beta} + \boldsymbol{Z'_{ij}b_i}\right\} \\ g\left(E\left(y_{ij}|\boldsymbol{b_i}\right)\right) &= g\left(\mu_{ij}\right) = \eta_{ij} = \boldsymbol{X'_{ij}\beta} + \boldsymbol{Z'_{ij}b_i} \end{aligned} \tag{4.1}$$

and the conditional variance is given by:

$$var\left(Y_{ij}\middle|\boldsymbol{b_i}\right) = \emptyset V\left(\mu_{ij}\right),$$

where $V(\mu_{ij})$ is the variance function that depends on μ_{ij} and thus η_{ij}, which is the linear predictor.

In the conditional variance equation above, Y_{ij} is the distribution of the exponential family that can be assumed to be conditional on the random effects $\boldsymbol{b_i}$, and the link function $g(\cdot)$, which is applied to the conditional mean of Y_{ij} given $\boldsymbol{b_i}$ to obtain the conditional linear predictor, and, finally, the linear predictor consisting of the fixed effects, $\boldsymbol{X'_{ij}\beta}$ and random effects, $\boldsymbol{Z'_{ij}b_i}$ whereby the distribution is assigned to $\boldsymbol{b_i}$.

Furthermore, $c(\,)$ and $d(\,)$ are known functions, $\emptyset$ is the dispersion parameter, and θ_{ij} is the canonical parameter that is defined implicitly by the mean of Y_{ij}, that is μ_{ij}, conditional on $\boldsymbol{b_i}$ Similar to the LME, the random effects, $\boldsymbol{b_i}$, are assumed to follow a normal distribution.

4.3.2 Marginal properties

As the model formulation of GLMM in Equation 4.1 is conditional on $\boldsymbol{b_i}$, the marginal properties of Y_{ij} can now be derived. McCulloch and Searle (McCulloch, 2003; McCulloch, 2001) have done this in great detail. The mean and variance of the marginal properties are briefly outlined. Furthermore, examples of marginal properties when the log link function is

used in the GLMM are provided; i.e., the function $g() = \log()$ and thus $g^{-1}() = \exp()$. This choice of link function is used when the response variable (Y_{ij}) follows the Poisson distribution.

4.3.2.1 Mean of Y_{ij}

Iterated expectations can be used to determine the mean of Y_{ij} as follows:

$$E(Y_{ij}) = E\{E(Y_{ij}|\boldsymbol{b_i})\}$$
$$= E\{g^{-1}(\boldsymbol{X'_{ij}\beta} + \boldsymbol{Z'_{ij}b_i})\}.$$

As an example, assuming that the log link function is used with a single random effect for each individual, $b_i \sim N(0, \sigma_b^2)$, so that the linear predictor is given by:

$$\eta_{ij} = \boldsymbol{X'_{ij}\beta} + b_i$$

Then, the marginal mean is given by:

$$E(Y_{ij}) = exp\left\{\boldsymbol{X'_{ij}\beta} + ({}^{\sigma_b^2}/_2)\right\} = \mu_{ij}. \quad (4.2)$$

This derivation involves the use of the moment generating function (m.g.f.) of the normal distribution (McCulloch, 2003; McCulloch, 2001).

4.3.2.2 Variance of Y_{ij}

Given that the variance of $Y_{ij} \mid \boldsymbol{b_i}$ is equal to $V(\mu_{ij})$, where μ_{ij} is the conditional mean of $Y_{ij} \mid \boldsymbol{b_i}$, the marginal variance is given by:

$$var(Y_{ij}) = var\{E(Y_{ij}|\,\boldsymbol{b_i})\} + E\{var(Y_{ij}|\,\boldsymbol{b}_i)\}$$
$$= var\{g^{-1}(\boldsymbol{X'_{ij}\beta} + \boldsymbol{Z'_{ij}b_i})\} + E\{V[g^{-1}(\boldsymbol{X'_{ij}\beta} + \boldsymbol{Z'_{ij}b_i})]\}.$$

For example, assuming that the log link function is used with a single random effect for each individual, $b_i \sim N(0, \sigma_b^2)$ and Y_{ij} follows a Poisson distribution with mean μ_{ij}, which indicates the conditional variance of $Y_{ij} \mid b_i$, the linear predictor is given as follows:

$$\eta_{ij} = \boldsymbol{X'_{ij}\beta} + b_i.$$

The result is the marginal variance, which is given by:

$$var(Y_{ij}) = E(Y_{ij})\left\{exp(X'_{ij}\beta)\left[exp\left(3\sigma_b^2/2\right) - exp\left(\sigma_b^2/2\right)\right] + 1\right\}$$
$$= \mu_{ij} + \mu_{ij}^2[\exp(\sigma_b^2) - 1]. \quad (4.3)$$

It can be seen from Equation 4.3 that the marginal variance of Y_{ij} will always be greater than the mean as the expression $(\exp(\sigma_b^2) - 1)$ will always be greater than 1. This shows that although $Y_{ij} \mid b_i$ follows a Poisson distribution, the marginal distribution will not. As a result, the marginal distribution will be overdispersed, thereby highlighting how random effects can be used to model overdispersion (McCulloch, 2003; McCulloch, 2001).

4.3.3 Estimation

This section explains the three most commonly used estimation methods in GLMM.

4.3.3.1 Maximum likelihood estimation (MLE)

Parameter estimation in a GLMM is determined by the maximisation of marginal likelihoods after integrating out random effects. The marginal likelihood contribution for the i^{th}subject is found by integration over $\boldsymbol{b_i}$, for $i = 1, \dots, q$, with each subject having n_i observations such that the total number of observations $n = \sum_{i=1}^{q} n_i$.

The marginal likelihood contribution is thus given as:

$$L_i(\boldsymbol{\beta}, \boldsymbol{D}, \emptyset) = \int \prod_{j=1}^{n_i} f_{ij}\left(y_{ij}\middle|\boldsymbol{b_i}, \boldsymbol{\beta}, \emptyset\right) f(\boldsymbol{b_i}|\boldsymbol{D})\, d\boldsymbol{b_i}. \quad (4.4)$$

This contribution is used to obtain the likelihood for ***β, D,*** *and* ∅*,* which is:

$$L(\boldsymbol{\beta}, \boldsymbol{D}, \emptyset) = \prod_{j=1}^{q} f_i(y_i|\boldsymbol{\beta}, \boldsymbol{D}, \emptyset) = \prod_{j=1}^{q} \int \prod_{j=1}^{n_i} f_{ij}\left(y_{ij}\middle|\boldsymbol{b_i}, \boldsymbol{\beta}, \emptyset\right) f(b_i|\boldsymbol{D})\, db_i. \quad (4.5)$$

As this integration is usually analytically intractable, various algorithms have been developed to solve it. The Gauss-Hermite quadrature (GHQ) is a numerical method of approximating the integral. This approach evaluates only single integrals if the random effects are independent. It can also not be used when high dimensional integrals are needed, such as when measurement data have subjects designed as cross-random effects (Lee et al., 2006). It is possible to overcome the shortcomings of the GHQ by using simulation methods. McCulloch (1994) developed a framework for ML and restricted MLE of variance components from binary data

using the Monte Carlo Expectation-Maximisation (EM) method. Other examples include Gibbs sampling and the simulated ML method (Karim & Zeger, 1992; McCulloch, 1997). It is worth noting that simulation methods have shortcomings of being computationally intensive and inaccurate in parameter estimation (Hobert & Casella, 1996). Approximation methods are another method of estimation. The algorithm proposed by Schall (Schall, 1991) yields approximate ML or quasi-ML estimates for the fixed effects and dispersion components and approximate empirical Bayes estimates for the random effects.

4.3.3.2 Empirical Bayes estimation

The empirical Bayes method provides the best linear unbiased predictions. Detecting subjects or groups of subjects that evolved differently over time is made more accessible by using random effects, which reflect between-subject variability. The predictions are also important when the study aims to predict subject-specific evolutions. Random effects are predicted based on the posterior distribution with the probability density function given as follows:

$$f_i\left(\boldsymbol{b_i}|y_{ij},\boldsymbol{\beta},\boldsymbol{D},\emptyset\right)=\frac{f_i\left(y_{ij}\middle|\,\boldsymbol{b_i},\boldsymbol{\beta},\emptyset\right)f(\boldsymbol{b_i}|\,\boldsymbol{D})}{\int f_i\left(y_{ij}\middle|\,\boldsymbol{b_i},\boldsymbol{\beta},\emptyset\right)f(\boldsymbol{b_i}|\,\boldsymbol{D})\,d\boldsymbol{b_i}}. \tag{4.6}$$

Since this posterior density usually does not have a normal distribution, the posterior mode is used as a point predictor of $\boldsymbol{b_i}$ instead of the posterior mean. Therefore, the predictor $\widehat{\boldsymbol{b}}_{\boldsymbol{i}}$ maximises $f_i(\boldsymbol{b_i}|y_{ij},\boldsymbol{\beta},\boldsymbol{D},\emptyset)$ with an ML estimate replacing the unknown parameters.

4.3.3.3 Bayesian methods

Estimating parameters using the Bayesian approach is widely used by researchers. Methods such as these assign prior distributions to $f(\boldsymbol{\beta})$, $\emptyset$, and **D** assuming independence among them. As a result, the prior density function for $f(\boldsymbol{\beta})$ typically follows a normal distribution or a flat, non-informative prior (Molenberghs & Verbeke, 2005).

Jeffreys priors can be used to calculate the **D** and $\emptyset$ prior densities, respectively, denoted as $f(\mathbf{D})$ and $f(\emptyset)$ (Gelman et al., 1995). Despite this, such a choice of priors can lead to incorrect posteriors (Fahrmeir et al., 1994). In the literature, an alternative approach was proposed by several authors, where they used proper but highly dispersed inverted Wishart (IW) priors for **D**, such that $\mathbf{D}\sim IW(\xi,\varphi)$, where ξ and φ are hyper-parameters that must be selected carefully.

The posterior distribution can be expressed as follows after the prior distributions have been specified:

$$f(\boldsymbol{\beta}, \mathbf{D}, \phi, \boldsymbol{b_1}, \dots, \boldsymbol{b_q}) \propto \prod_{i=1}^{q} \prod_{j=1}^{n_i} f_i(y_{ij} | \boldsymbol{\beta}, \emptyset, \boldsymbol{b_i}) f(\boldsymbol{b_i}|\mathbf{D}) f(\mathbf{D}) f(\boldsymbol{\beta}) f(\emptyset).$$

Gibbs sampling is one of the standard algorithms for estimating the fixed and random effects based on samples drawn from the posterior distribution (Karim & Zeger, 1992).

4.3.4 Inference

The inferences for the estimated parameters of GLMMs are based on classical ML theory since ML is used to fit GLMMs. Consequently, once the appropriate model has been fitted, the estimators of the parameters have asymptotically normal distributions, with the true value of the parameter as the mean and the inverse Fisher information matrix as the covariance matrix. A Wald-type test can be performed on the estimates since they have an asymptotic normal distribution.

It is also possible to test composite hypotheses using a quadratic standardised Wald statistic compared to the chi-squared distribution (Molenberghs & Verbeke, 2005). Likelihood ratio and score tests are other tests that can be used. When interest is in the inference for the variance components in **D**, the classical Wald, likelihood ratio, and score tests can be used as long as the hypotheses being tested are not on the boundary of the parameter space. For example, when a researcher wants to test whether the variance σ_b^2 of a random effect is equal to 0; that is $H_0: \sigma_b^2 = 0$ versus $H_1: \sigma_b^2 > 0$. In such a case, the Wald, likelihood ratio, and score tests cannot be used because the regularity conditions have not been met (Stram & Lee, 1995). Several authors have covered the subject of testing hypotheses on the boundary of parameter space, including Self and Liang (Self & Liang, 1987).

4.4 THE EXTENDED COX MODEL

Section 3.5 described the extended Cox model as an extension of the Cox PH regression model used to fit exogenous time-dependent covariates using the counting process formulation (Anderson & Gill, 1982).

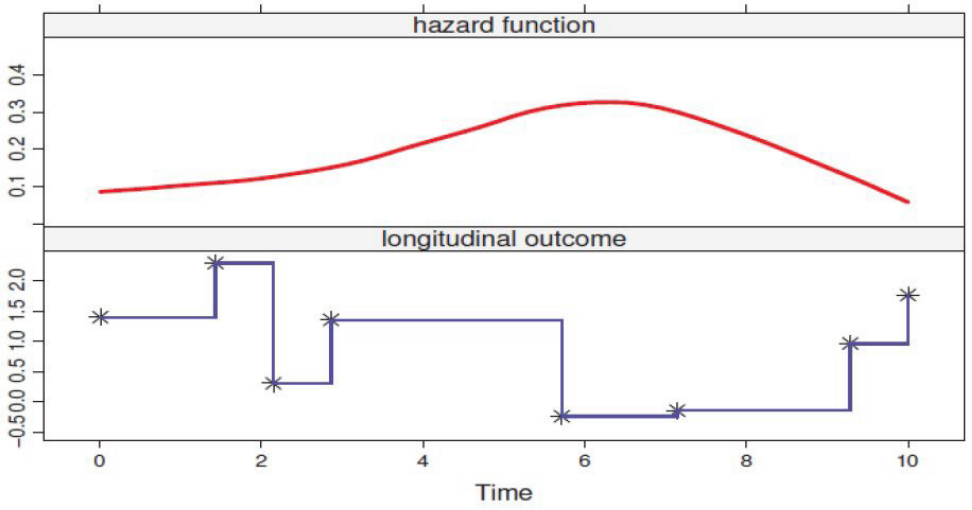

Figure 4.1: A visual representation of the extended Cox model (Rizopoulos, 2012**)**

Figure 4.1 shows a broader picture of how the time-dependent covariates are handled in the extended Cox model framework. A time-dependent covariate can either be endogenous (internal) or exogenous (external). Endogenous covariates are those whose values are affected by the probability and timing of the event. It is only possible to measure these covariates when the patient is alive. In the IMPI dataset, variables such as white blood cell count and systolic blood pressure are good examples. The exogenous covariates, however, are not dependent on patients' survival. It is important to note that the values do not change based on when and if the event occurred. Such a variable is one whose value can be predicted in advance at a future date. An example would be a patient's age, as once it is known at the beginning of the study, it will also be known in the future.

The extended Cox model (also known as the Anderson and Gill model), where the longitudinal marker was assumed to remain constant between visits, is expressed as:

$$h_i(t\ |Y_i(t), x_i) = h_0(t)R_i(t)\ exp\{\boldsymbol{X_i\beta'} + \alpha Y_i(t)\}, \tag{4.7}$$

where $h_0(t)$ is the baseline hazard function, X_i is a vector of baseline covariates, $Y_i(t)$ are time-dependent covariates, $R_i(t)$ is the risk set, and β_i & α_j are the regression coefficient vectors.

Accordingly, HRs cannot be assumed to be constant in Equation 4.7, since the covariates are time-dependent.

The regression coefficient vectors (β_i & α_j) have the same interpretation and are estimated based on the partial log-likelihood expressed as (Rizopoulos, 2012):

$$p\ell(\boldsymbol{\beta}, \alpha) = \sum_{i=1}^{n} \int_0^{\infty} \{R_i(t)\ exp\{\boldsymbol{X_i\beta'} + \alpha Y_i(t)\} - log\left[\sum_j \{R_i(t)\ exp\{\boldsymbol{X_i}\beta' + \alpha Y_i(t)\}\right]\} dN_i(t),$$

where, R_i represents the risk set, N_i represents the counting process, β represents the baseline covariates, and α represents the time-varying risk coefficient. Specifically, alpha or the association parameter is interpreted as an increase in the risk of an event (death or constriction) at a time t caused by a one-unit decrease in a time-varying covariate (ECG characteristic variables) at the same time point.

4.5 JOINT MODELS FOR LONGITUDINAL AND TIME-TO-EVENT DATA

The notation by Rizopoulos et al. (2012) is adopted in this section. As discussed in Section 4.4, using the extended Cox model is appropriate when the covariates are exogenous time-dependent. As a result, it does not appropriately handle endogenous time-varying covariates such as biomarkers since the underlying assumption that covariates change only during these measurements is unrealistic. Joint modelling is therefore a more appropriate technique when assessing the association between an endogenous time-varying covariate and survival. A key concept behind the joint model is that it couples the survival model, which is of primary interest, with a suitable method of repeated measurements of the covariate that can explain its specific features. The following techniques can be used to model data with a time-dependent covariate: the extended Cox model, a two-stage joint model, or joint models using shared parameters.

This study used the joint model framework to analyse the IMPI trial dataset in order to determine patient survival. The study first examined the two-stage joint model approach since the extended Cox model was covered in Section 4.4. In the next step, the sub-models of the joint model were specified and how the parameters and random effects can be estimated using MLE and Bayesian estimation respectively was discussed. This was concluded by looking at the parameterisation of the joint model (associations between the longitudinal outcome and the hazard of subsequent events) and the multivariate joint models.

4.5.1 A two-stage approach to the joint model

When interest is in the evolution of the longitudinal responses and time-to-event data, the simplest method is to fit an extended Cox model with the longitudinal measurements specified as time-dependent covariates. However, the shortcoming of this model is that it underestimates the true association between the hazard and underlying response because of measurement errors

that yield biased estimates (Prentice, 1982). It is for this reason that a two-stage joint modelling has been proposed as one of the classical approaches in joint modelling. The model is applied in two steps as follows:

In the first step, a linear mixed model is fitted to the data, and the fitted linear mixed model is given as follows:

$$Y_i = X_i\boldsymbol{\beta} + Z_i\boldsymbol{b_i} + \boldsymbol{\varepsilon_i}, \tag{4.8}$$

where a vector Y_i represents the observed outcomes for the subject, b_i are random effects, which are assumed to be distributed as $b_i \sim N(0, \mathrm{G})$; G represents the covariance matrix for the random effects and residuals; $\boldsymbol{\varepsilon_i} \sim N(\mathbf{0}, \mathbf{R})$, where $R = \sigma^2\mathbf{I_i}$; $\boldsymbol{X_i}$ and Z_i are design matrices for both the fixed effects and random effects and the usual assumption is that $R = \sigma^2\mathbf{I_i}$.

In the second step, subject-specific predictions from the fitted linear mixed model are extracted (4.8), and the extracted predictions in the extended survival model (4.9) are used as time-varying covariates. The survival model is given by:

$$h_i(t\,|m_i(t)) = h_0(t)\,exp\{\boldsymbol{X_i'\tau} + \widehat{m}_i(t)\}, \tag{4.9}$$

where $\widehat{m}_i(t) = \{m_i(s), 0 \leq s < t\}$ is the longitudinal history of the i^{th} subject (Y_i) and X_i is the fixed effects matrix, which are predominantly involved in the second step, and $\boldsymbol{\tau}$ represents a vector of coefficients for the fixed effects.

In the two-stage model, the errors from the first step are not carried through to the second step, although Mauff et al. (2020) addressed this using importance sampling. This therefore gives the hope that the model would be widely used because it is simple and not computer intensive.

4.5.2 Shared parameter joint model

This section adopts and introduces the notation from Rizopoulos et al. (2012). Let $\mathrm{T}_i^\star$ be the true event time for the i^{th} subject, T_i the observed event time, δ_i the event indicator that equals 1 if the event occurred and 0 otherwise, and $y_i(t)$ the observed longitudinal measurements for the i^{th} subject at time point t. The shared parameter joint model for longitudinal and time-to-event data uses all repeated measurements, accounts for measurement error, reduces bias, and maximises efficiency. In this class of joint model, the random effects are shared between the

longitudinal and survival processes (Tsiatis & Davidian, 2004). In addition, two sub-models make up the joint model: longitudinal and survival.

There are three steps in joint modelling, following the adopted notation:

Step 1: Assume that the term $m_i(t)$ is the known and true unobserved longitudinal marker value at time t. The standard relative risk model is therefore defined as follows:

$$h_i(t\,|\boldsymbol{\mathcal{M}_i}(t)) = h_0(t)exp\{\gamma'\omega_i + \alpha m_i(t)\}, \tag{4.10}$$

where $\mathcal{M}_i(t) = \{m_i(s), 0 \le s < t\}$ is the longitudinal history, α measures the strength of association between the marker and the risk for an event, $\boldsymbol{\omega_i}$ are the baseline covariates, and γ' represents a vector of regression coefficients.

Step 2: The covariate history for each subject is constructed based on the observed longitudinal responses $(y_i(t))$:

$$y_i(t) = m_i(t) + \boldsymbol{\varepsilon_i}(t)$$
$$= \boldsymbol{x_i^T}(t)\boldsymbol{\beta} + \boldsymbol{z_i^T}(t)\boldsymbol{b_i} + \boldsymbol{\epsilon_i}(t), \qquad \boldsymbol{\epsilon_i}(t) \sim N(0, \sigma^2), \quad \boldsymbol{b_i} \sim N(0, \boldsymbol{D})$$

where $\boldsymbol{x_i^T}(t)$ and $\boldsymbol{\beta}$ are the fixed effects of the model, and $\boldsymbol{z_i^T}(t)$ and $\boldsymbol{b_i}$ are the random effects of the model.

Step 3: The third step describes a conditional joint distribution relating to the two processes discussed in Steps 1 and 2 (Tsiatis & Davidian, 2004):

$$p(\boldsymbol{y_i}, \mathrm{T}_i, \delta_i) = \int p(\boldsymbol{y_i} \mid \boldsymbol{b_i})\left\{h(\mathrm{T}_i \mid \boldsymbol{b_i})^{\delta_i} S(\mathrm{T}_i \mid \boldsymbol{b_i})\right\} p(\boldsymbol{b_i})\ d\boldsymbol{b_i},$$

where $\boldsymbol{b_i}$ represents a vector of random effects that explains the i^{th} subject deviation from the population, $p(.)$ represents a density function, and $S(.)$ represents a survival function.

The main assumption of the joint model is that the longitudinal outcome is independent of the time-to-event outcome and the repeated measurements in the longitudinal outcome are independent of each other:

$$p(y_i, \mathrm{T}_i, \delta_i \mid b_i) = p(y_i \mid b_i)p(\mathrm{T}_i, \delta_i \mid b_i)$$
$$p(y_i \mid b_i) = \prod_j p(y_{ij} \mid b_i)$$

A longitudinal marker history determines the survival function. A survival function, which is an important component of the likelihood function, is determined by the whole longitudinal history:

$$S_i(t\,|b_i) = exp\left(-\int_0^t h_0(s)\,exp\{\gamma^{\mathrm{T}}\omega_i + \alpha m_i(s)\}\,ds\right).$$

4.5.3 Parameter estimation for the joint model for longitudinal and time-to-event data

A semi-parametric ML is the standard estimation method for parameters in the joint model framework (Henderson et al., 2000; Hsieh et al., 2006; Tsiatis & Davidian, 2001). The method is performed by maximising the log-likelihood function corresponding to the joint distribution of the survival and longitudinal outcomes $\{\mathbf{T_i}, \boldsymbol{\delta_i}, \boldsymbol{Y_i}\}$. The joint conditional distribution of the i^{th} subject is then presented as follows (Rizopoulos, 2012; Tsiatis & Davidian, 2004):

$$p(\boldsymbol{Y_i}, \mathbf{T_i}, \boldsymbol{\delta_i}|\, \boldsymbol{b_i}; \boldsymbol{\theta}) = p(\boldsymbol{Y_i}, \mathbf{T_i}, \boldsymbol{\delta_i}|\, \boldsymbol{b_i}; \boldsymbol{\theta}) p(\boldsymbol{Y_i}\mid \boldsymbol{b_i}; \boldsymbol{\theta})$$

and

$$p(\boldsymbol{Y_i}, \mathbf{T_i}, \delta_i|\, \boldsymbol{b_i}; \boldsymbol{\theta}) = \prod_{j=1}^{n_i} p\{\boldsymbol{Y_i}(\mathrm{T}_{ij}|\boldsymbol{b_i}; \boldsymbol{\theta})\}$$

where $\boldsymbol{\theta} = (\theta_T', \theta_Y', \theta_b')$ refers to a full parameter vector that consists of $\boldsymbol{\theta}_T$ denoting the parameters for the event time outcome, $\boldsymbol{\theta}_Y$ denotes the parameters for the longitudinal outcomes, and $\boldsymbol{\theta}_b$ denotes the parameters of the random effects covariance matrix. It assumes that $\boldsymbol{b_i}$, the vector of random effects, are shared by both the longitudinal and survival processes. The log-likelihood contributing to the i^{th} subject is as follows:

$$\begin{aligned}\log\, p(\mathbf{T}_i, \mathbf{Y}_i, \boldsymbol{\delta}_i; \boldsymbol{\theta}) &= \log \int p(\mathbf{T}_i, \boldsymbol{\delta}_i, \mathbf{Y}_i, \mathbf{b}_i; \boldsymbol{\theta}) d\mathbf{b}_i \\ &= \log \int p(\mathbf{T}_i, \boldsymbol{\delta}_i, \mid \mathbf{b}_i; \boldsymbol{\theta}, \boldsymbol{\beta}) \left[\prod_{j=1}^{n_i} p\{\mathbf{Y}_i(T_{ij}) \mid \mathbf{b}_i; \boldsymbol{\theta}_b\}\right] p\left(\mathbf{b}_i; \overline{\boldsymbol{\theta}}_b\right) d\mathbf{b}_i. \quad (4.11)\end{aligned}$$

Equation (4.11) further assumes that the censoring mechanism and the visiting process are independent of the true event times and future longitudinal measurements, given the observed history. The integral is intractable because of the random effects; numerical techniques are therefore applied, e.g. EM algorithm (Lipsitz et al., 2002). Therefore, for any time point *t*, the observed history is all the available information for the longitudinal process prior to t.

The density function of the time-to-event process is presented below:

$$\begin{aligned} p(T_i, \delta_i \mid b_i; \theta_t, \beta) &= h_i(T_i \mid \boldsymbol{\mathcal{M}_i}(T_i); \theta_t, \beta)^{\delta_i} \boldsymbol{S_i}(T_i \mid \boldsymbol{\mathcal{M}_i}(T_i); \theta_t, \beta) \\ &= [h_0(T_i)\exp\{\gamma^T w_i + \alpha m_i(T_i)\}]^{\delta_i} \\ &\times \exp\left(-\int_0^{T_i} h_0(s)\exp\{\gamma^T w_i + \alpha m_i(s)\}ds\right), \end{aligned}$$

where $h_0(\cdot)$, the hazard function, can assume any underlying distribution of time. The log-likelihood function $\ell(\boldsymbol{\theta}) = \sum_i \log p(\boldsymbol{T_i}, \boldsymbol{Y_i}, \boldsymbol{\delta_i}; \boldsymbol{\theta})$ can be maximised with respect to $\boldsymbol{\theta}$ by using standard optimisation algorithms such as the EM by Dempster et al. (1977) or the Newton-Raphson algorithm by Hunter and Lange (2004). In a joint modelling framework, the EM algorithm is mainly used (Rizopoulos, 2012).

The joint density function of the longitudinal process including the random effects is presented in equation (4.12):

$$\begin{aligned} p(y_i \mid b_i; \theta)p(b_i; \theta) &= \prod_j p\{y_i(t_{ij}) \mid b_i; \theta_y\}p(b_i; \theta_b) \\ &= (2\pi\sigma^2)^{n_i/2}\exp\left\{\|y_i - \mathbf{X}_i\beta - Z_i b_i\|^2/2\sigma^2\right\}, \\ &\times (2\pi)^{qb/2}\det(D)^{-1/2}\exp(-b_i^T D^{-1} b_i/2) \end{aligned} \tag{4.12}$$

where q_b is the dimension of the random effects vector, and $||x|| = \{\sum_i x_i^2\}^{1/2}$ is the norm of the Euclidean vector.

The random effects, b_i, are estimated using the Bayesian methods proposed by Rizopoulos (Rizopoulos, 2012). The idea behind the method is that estimation is performed based on the posterior distribution. Given the prior distribution $p(b_i; \theta)$ and a conditional likelihood function $p(b_i|Y_i, T_i, \delta_i, y_i; \theta)$, then:

$$\begin{aligned} p(b_i \mid T_i, \delta_i, y_i; \theta) &= \frac{p(T_i, \delta_i \mid b_i; \theta)p(y_i \mid b_i; \theta)p(b_i; \theta)}{p(T_i, \delta_i, y_i; \theta)}. \\ &\propto p(T_i, \delta_i \mid b_i; \theta)p(y_i \mid b_i; \theta)p(b_i; \theta) \end{aligned} \tag{4.13}$$

The equation (4.13) has no closed solution, so it is solved numerically and the measures of location used are the mean and the mode, as shown below:

$$\begin{cases} \bar{b}_i \int b_i p(b_i \mid Y_i, \delta_i, y_i; \theta) \text{ and} \\ \hat{b}_i \ \arg max_b \{\log\ p(b_i \mid Y_i, \delta_i, y_i; \theta)\} \end{cases}.$$

4.6 PARAMETRISATIONS OF THE JOINT MODEL

This section examines various parameterisations of the joint model introduced in the previous section. Specifically, the focus is on parameterisations for the association between the time-to-event outcome and longitudinal processes involving exogenous time-dependent covariates. The R packages used were JM, JMBayes, and JMBayes2.

Equation 4.10 assumes a current value parameterisation. The parameter α captures the strength of the association between the longitudinal and the event processes. However, this assumption may be violated and lead to inappropriate conclusions from the results obtained in medical research and other disciplines. For this reason, there are assumptions to express the association structure between the risk of an event and longitudinal outcome. In this setting, the longitudinal and event outcomes are linked using the following assumptions: the lagged effect parameterisations, the interaction effect parameterisations, the cumulative effect parameterisations, the weighted cumulative effect parameterisations, and the time-dependent slopes parameterisations.

In the lagged effect parameterisations, the risk of an event was assumed to be associated with the longitudinal marker level at a proceeding time point:

$$h_i(t) = h_0(t)\exp[\gamma^T w_i + \alpha m_i\{\max(t-c,0)\}],$$

where *c* represents the time lag one is interested in.

Under the interaction effect parameterisation, the biomarker was allowed to be different for different subgroups of the intended population. This is done by including an interaction term of the linear predictor with baseline covariates. The formulation is shown below:

$$h_i(t) = h_0(t)\exp\left[\gamma^T w_{i1} + \alpha^T\{w_{i2} \times m_i(t)\}\right]$$

Where, ω_{i1} represents a vector that accommodates the direct effect of the baseline covariates, and ω_{i2} contains covariates that interact with $m_i(t)$ by linking different associations for different subgroups of the data (Rizopoulos, 2012).

The time-dependent slope parameterisation is the third extension that allows the risk of an event to depend on other features of the longitudinal trajectory since it is known that each subject has a unique slope followed up at a different time.

It is therefore assumed that the risk of an event depends on the current value of the longitudinal trajectory in the formulation below:

$$h_i(t) = h_0(t)\exp\{\gamma^T w_i + \alpha_1 m_i(t) + \alpha_2 m_i'(t)\},$$

where:

$$m_i'(t) = \frac{d}{dt}m_i(t) = \frac{d}{dt}\{x_i^T(t)\beta + z_i^T(t)b_i\}$$

α_1 measures the degree of the association between the risk of an event at time t and the current longitudinal outcome at the same time. On the other hand, α_2 measures the degree of the association between the risk of an event at time t and the slope of the longitudinal outcome at the same time provided $m_i(t)$ is constant.

The cumulative effect parameterisation can be viewed as an extension of the time-dependent slope parameterisation. The difference is that the risk of an event in cumulative effect parameterisations depends on the whole longitudinal trajectory history. This is performed by including the linear predictor of the relative risk sub-model as the integral of the longitudinal trajectory, representing the cumulative effect of the longitudinal outcome up to the time point t. In addition, this approach will allow an association of the longitudinal history with the hazard of the event. The formulation is shown below:

$$h_i(t) = h_0(t)\exp\left\{\gamma^T w_i + \alpha\int_0^t m_i(s)ds\right\},$$

where α measures the strength of the association between the risk of an event and the cumulative longitudinal trajectory at a particular time t.

Finally, the random effects parameterisation only includes the random effects of the longitudinal sub-model in the linear predictor of the relative risk model as follows:

$$h_i(t) = h_0(t)\exp(\gamma^\top w_i + \alpha^\top b_i),$$

where α is a vector of association parameters, each measuring the association between the random effect and the hazard for an event. In this case, the simple random intercepts and random slopes were assumed. Furthermore, the random effects express subject-specific departures from the average intercept and average slope (Rizopoulos, 2012). A major focus of this study was on the current value parameterisation of the association structure, which is based

on the dataset (IMPI), which shows how ECG characteristics influence the risk of constrictive pericarditis and death, or death and constrictive pericarditis risk, respectively.

4.7 EXTENDED COX MODEL FOR NON-NORMAL TIME-DEPENDENT VARIABLES

Section 3.5 described the extended Cox model as an extension of the Cox PH regression model that is used to fit exogenous time-dependent covariates using the counting process formulation (Anderson & Gill, 1982). Section 4.4 discussed in detail the implementation and the definition of how the time-dependent covariates are handled in the extended Cox model setting. However, there are cases where the time-dependent covariates are not normal. In such a case, the $Y_i(t)$, which are the time-dependent covariates, are assumed to be binary in Equation 4.7. The ECG characteristics, which are the time-dependent covariates from the IMPI trial dataset presented in this study, are binary-coded as either 0 or 1. For example, ECG covariates are time-dependent, whereby at a particular visit, it is recorded as whether the patient has a normal or abnormal ECG rate. The study therefore compared those who experienced the event at time t; i.e., patients who have a normal measurement, with who that did not experience it; i.e., who had an abnormal measurement coded as 1.

4.8 JOINT MODEL FOR NON-NORMAL LONGITUDINAL DATA

So far, joint models with a continuous longitudinal outcome have been discussed. There are, however, situations in which the longitudinal outcomes can be binary or categorical. The standard linear models are extended in this setting using the GLMM discussed in Section 4.3. More specifically, the model is developed to handle non-normal longitudinal data. These models are fitted using Bayesian techniques based on Dirichlet priors to estimate the parameters and this is what was focused on in this study. As all the analyses were implemented in R, the latest version of JMBayes2 enabled fitting the joint models in real computing time.

4.9 NON-NORMAL JOINT MODEL

The joint model structure for the non-normal longitudinal outcome and time-to-event data is first defined in detail [90]. The sample from the target population is $D_n = \{\mathrm{T}_i, \mathrm{T}_i^U, \delta_i, Y_i; i = 1, \dots, n\}$, where T_i^* represents the true event time for the i^{th} subject, T_i and T_i^U are observation times of participants, and $\delta_i \in 0, 1, 2, 3$ is the event indicator, with 0 corresponding to right

censoring ($T_i^* > T_i$), 1 to a true event ($T_i^* = T_i$), 2 to left censoring ($T_i^* < T_i$), and 3 to interval censoring $\left(T_i < T_i^* < T_i^U\right)$. y_{ik} is defined as the n_{ik} X 1 longitudinal vector response for the K longitudinal outcome for each subject, with element Y_{ijk} representing the value of that longitudinal outcome at time point $t_{ijk}, j = 1, \dots, n_{ik}$. The multivariate GLMM is adopted in a joint model framework to accommodate the different longitudinal outcomes (McCulloch, 2003).

In particular, it was assumed that the conditional distribution of the outcome given the vector of random effects b_i is the member of the exponential family with the linear predictor as follows:

$$g\{E(Y_i(t) \mid b_i)\} = \eta_i(t) = X_i^\top(t)\beta + Z_i^\top(t)b_i, \mathbf{b}_i^T \sim N(0, D)$$

where $g(\cdot)$ represents the known one-to-one monotonic link function, $Y_i(t)$ represents the value of the longitudinal outcome for i^{th} subject at time t, $X_i^\top(t)$, and $Z_i^\top(t)$ represents the fixed and random effects vectors respectively. Furthermore, it was assumed that the vector of random effects $b_i = (b_{1i}, b_{2i}, \dots b_{in})^\mathrm{T}$ follows a normal distribution with a mean of zero and the variance/covariance matrix D. For the survival process, the hazard model for the longitudinal outcome is given as follows:

$$\begin{aligned} h_i(t) &= h_0(t)\exp\left\{\gamma^T\mathbf{X}_i(t) + \{\boldsymbol{\mathcal{H}_i}(t), X_i(t), b_i, \alpha\}\right\} \\ &= h_0(t)\exp\left\{\gamma^T\mathbf{X}_i(t) + \alpha\eta_i(t)\right\} \end{aligned}$$

where $\boldsymbol{\mathcal{H}_i}(t) = \{\eta_i(s), 0 \le s < t\}$ represents the longitudinal process history up to t, $\mathbf{X}_i(t)$ represents the vector of covariates with its corresponding regression coefficient γ, and $h_0(t)$ represents the baseline hazard function modelled using the B-splines approach as follows:

$$\log\, h_0(t) = \sum_{q=1}^{Q} \gamma_{h_{0,q}} B_q(t, v)$$

where $B_q(t, v)$ represents the q^{th} basis function of a B-spline with knots $v_1, \dots, v_Q$, and γ_{h0} represents the spline coefficients vector. In order to avoid worrying if the knots are placed correctly, adding a large number of knots is beneficial. In terms of smoothness, the B-spline regression coefficients γ_{h0} were penalised using differences in penalty.

Joint models for longitudinal responses also use different parametrisation structures for association parameters, just like the standard joint model, as follows:

$$f\{\boldsymbol{\mathcal{H}_i}(t), w_i(t), b_i, \alpha\}\& = \alpha\eta_i(t) \tag{4.13a}$$

$$f\{\boldsymbol{\mathcal{H}_i}(t), w_i(t), b_i, \alpha\}\& = \alpha_1\eta_i(t) + \alpha_2\eta_i'(t) \text{ with } \eta_i'(t) = \frac{d\eta_i(t)}{dt} \tag{4.13b}$$

$$f\{\boldsymbol{\mathcal{H}_i}(t), w_i(t), b_i, \alpha\}\& = \alpha \int_0^t \eta_i(s)ds \tag{4.13c}$$

$f\{\boldsymbol{\mathcal{H}_i}(t), w_i(t), b_i, \alpha\}\& = \alpha\eta_i(t)$ in Equation 4.13a, it was assumed that the risk of an event at a particular time t may be associated with the longitudinal outcome at the same time point. Furthermore, for both Equations 4.13b and 4.13c, it was assumed that the risk of an event at a particular time t is associated with the slope of the longitudinal outcome at the same time point or the cumulative trajectory up to time t respectively.

4.9.1 Parameter estimation

Parameter estimation in the joint model is done by using the Bayesian approach. Given the observed data, the posterior distribution of the model parameters is based on the assumptions that the longitudinal outcomes are independent of event times and that the longitudinal responses of the parameters are independent of each other as well. As a result, the posterior distribution would be as follows:

$$p(\theta, b \mid Y_i) \propto \prod_{i=1}^{n} \prod_{j=1}^{n_l} p\big(y_{ij} \mid b_i, \theta\big)p\big(T_i, T_i^U, \delta_i \mid b_i, \theta\big)p(b_i \mid \theta)p(\theta),$$

where θ represents the full parameter vector, which includes the fixed effects from both the sub-models, the association parameters from the survival sub-model, and the variance components from the GLMM.

Furthermore, $p\big(y_{kij} \mid b_{ki}, \theta\big)$ represents the conditional distribution for the longitudinal part of the model, which is represented as follows:

$$p\big(y_{kij} \mid b_{ki}, \theta\big) = \exp\left\{\frac{\big[y_{kij}\psi_{kij}(b_{ki}) - c_k\{\psi_{kij}(b_{ki})\}\big]}{a_k(\varphi)} - d_k\big(y_{kij}\varphi\big)\right\},$$

where $\psi_{kij}(b_{ki})$ and φ denote the natural and dispersion parameters in the exponential family respectively, and $c_k(\cdot), a_k(\cdot)$ and $d_k(\cdot)$ are known functions that specify the member of the exponential family.

Finally, $p\left(T_i T_i^U, \delta_i \mid b_{ki}\right)$ represents the density distribution for the survival part of the model, which is represented as follows:

$$p\left(T_i, T_i^U, \delta_i \mid b_{ki}, \theta\right) = \left\{h_i\left(T_i \mid \boldsymbol{\mathcal{H}_i}(T_i), w_i(T_i)\right)\right\}^{I(\delta_i=1)} \exp\left\{-\int_0^{T_t} h_i(s \mid \boldsymbol{\mathcal{H}_i}(s), w_i(s))ds\right\}$$
$$\times \left\{1 - \exp\left\{-\int_0^{T_t} h_i(s \mid \boldsymbol{\mathcal{H}_i}(s), w_i(s))ds\right\}\right\}^{I(\delta_i=2)}$$
$$\times \left\{\exp\left\{-\int_0^{T_t} h_i(s \mid \boldsymbol{\mathcal{H}_i}(s), w_i(s))ds\right\}\right.$$
$$\left. - \exp\left\{-\int_0^{T_t} h_i\left(s \mid \boldsymbol{\mathcal{H}_i}(s), w_i(s)\right)ds\right\}\right\}^{I(\delta_i=3)},$$

where $I(\cdot)$ represents the indicator function, the parameters T_i^U , $\delta_i = 2$, and $\delta_i = 3$ are used when the data are interval censored.

CHAPTER 5:
MULTIVARIATE JOINT MODELS FOR LONGITUDINAL AND TIME-TO-EVENT DATA

5.1 INTRODUCTION

This chapter introduces the multivariate linear mixed model, focusing on multiple Gaussian outcomes. This setting is extended to environments where longitudinal outcomes derive from a mix of exponential distributions using GLMMs. The posterior distribution functions associated with these models as part of the parameter estimation process were also written. A Bayesian approach was used to estimate the parameters of the multivariate joint model utilising MCMC and a Gibbs Sampler extension called JAGS (see Section 5.2). Estimates of all parameters from the model that were believed to have converged were obtained. These estimates included variance components from the GLMM, fixed effects, and association parameters from the survival sub-model. The model-fitting results follow in the next chapter.

So far, joint models with only one longitudinal outcome have been discussed. There are, however, situations in which multiple longitudinal outcomes can be used to predict the risk of an event. The standard linear models in this setting are extended using the GLMM discussed in Section 4.3. More specifically, the model was developed to handle multivariate longitudinal data. Due to the high number of random effects, the implementation of these models is computationally intensive, even though they are mathematically simple. These models are fitted using Bayesian techniques based on Dirichlet priors to estimate the parameters, which are what this study focused on. As all analyses were implemented in R, the latest version of JMBayes2 enables one to fit joint models in real computing time.

5.2 MULTIVARIATE JOINT MODEL

The study first defined in detail the multivariate joint model structure for multiple longitudinal outcomes and time-to-event data (Chi & Ibrahim, 2006; Hickey et al., 2016; Mauff et al., 2020) The sample from the target population was $D_n = \{\mathrm{T}_i, \mathrm{T}_i^U, \delta_i, Y_i; i = 1, \dots, n\}$, where T_i^* represents the true event time for the i^{th} subject, T_i and T_i^U are observation times of participants, and $\delta_i \in 0, 1, 2, 3$ is the event indicator, with 0 corresponding to right censoring ($\mathrm{T}_i^* > \mathrm{T}_i$), 1 to a true event ($\mathrm{T}_i^* = \mathrm{T}_i$), 2 to left censoring ($\mathrm{T}_i^* < \mathrm{T}_i$), and 3 to interval censoring

$\left(\mathrm{T}_i < \mathrm{T}_i^* < \mathrm{T}_i^U\right)$. y_{ik} is defined as the n_{ik} X 1 longitudinal vector response for the K longitudinal outcome for each subject, with element Y_{ijk} representing the value of that longitudinal outcome at time point $t_{ijk}, j = 1, \dots, n_{ik}$. In a joint model framework, the multivariate GLMM was adopted to accommodate the different longitudinal outcomes (McCulloch, 2003).

In particular, the study assumed that the conditional distribution of k^{th} outcome given the vector of random effects b_{ik} is the member of the exponential family with the linear predictor as follows:

$$g_{ik}\{E(Y_{ik}(t) \mid b_{ik})\} = \eta_{ki}(t) = X_{ik}^{\top}(t)\beta + Z_{ik}^{\top}(t)b_{ik}, \mathbf{b}_{ik}^{T} \sim MVN(0, D)\ , \tag{5.1}$$

where $g_{ik}(\cdot)$ represents the known one-to-one monotonic link function, Y_{ik} represents the value of the k^{th} longitudinal outcome for i^{th} subject at time t, $X_{ik}^{\top}(t)$, and $Z_{ik}^{\top}(t)$ represents the fixed and random effects vectors respectively. Furthermore, it was assumed that the vector of random effects $b_i = (b_{1i}, b_{2i}, \dots b_{in})^{\mathrm{T}}$ follows a multivariate normal distribution with a mean of 0 and the variance/covariance matrix D. For the survival process, the hazard model for the multiple longitudinal outcomes is given as follows:

$$\begin{aligned} h_i(t) &= h_0(t)\exp\left\{\gamma^T\mathbf{X}_i(t) + \sum_{k=1}^{K}\sum_{l=1}^{L_k} f_{lk}\{\mathcal{H}_{ik}(t), X_i(t), b_{ik}, \alpha_{lk}\}\right\} \\ &= h_0(t)\exp\left\{\gamma^T\mathbf{X}_i(t) + \sum_{k=1}^{K}\alpha_k\eta_{ki}(t)\right\}, \end{aligned}$$

where $\mathcal{H}_{ki}(t) = \{\eta_{ki}(s), 0 \le s < t\}$ represents the longitudinal process history up to t, $\mathbf{X}_i(t)$ represents the vector of covariates with its corresponding regression coefficient γ, and $h_0(t)$ represents the baseline hazard function modelled using the B-splines approach as follows:

$$\log h_0(t) = \sum_{q=1}^{Q}\gamma_{h_0,q}B_q(t, v),$$

where $B_q(t, v)$ represents the q^{th} basis function of a B-spline with knots $v_1, \dots, v_Q$ and γ_{h0} represents the spline coefficients vector. In order to avoid worrying if the knots are placed

correctly, adding a large number of knots is beneficial. In terms of smoothness, the B-spline regression coefficients γ_{h0} were penalised using differences in penalty.

Joint models for multiple longitudinal responses also use different parametrisation structures for association parameters, just like the standard joint model, as follows:

$$f\{\boldsymbol{\mathcal{H}_i}(t), w_i(t), b_i, \alpha\} = \alpha\eta_i(t) \tag{5.2a}$$

$$f\{\boldsymbol{\mathcal{H}_i}(t), w_i(t), b_i, \alpha\} = \alpha_1\eta_i(t) + \alpha_2\eta_i'(t) \text{ with } \eta_i'(t) = \frac{d\eta_i(t)}{dt} \tag{5.2b}$$

$$f\{\boldsymbol{\mathcal{H}_i}(t), w_i(t), b_i, \alpha\} = \alpha \int_0^t \eta_i(s)ds\ . \tag{5.2c}$$

In Equation 5.2a, it was assumed that the risk of an event at a particular time t may be associated with the longitudinal outcome at the same time point. Furthermore, in both Equations 5.2b and 5.2c, it was assumed that the risk of an event at a particular time t is associated with the slope of the longitudinal outcome at the same time point or the cumulative trajectory up to time t respectively.

5.2.1 Parameter estimation

Parameter estimation in the multivariate joint model is done using the Bayesian approach. Given the observed data, the posterior distribution of the model parameters was based on the assumptions that the longitudinal outcomes are independent of event times, that the multiple longitudinal outcomes are independent of each other, and that the longitudinal responses of the parameters are independent of each other as well. As a result, the posterior distribution would be as follows:

$$p(\theta, b \mid Y_{ki}) \propto \prod_{i=1}^{n} \prod_{k=1}^{K} \prod_{j=1}^{n_{kl}} p\big(y_{kij} \mid b_{ki}, \theta\big)p\big(T_i, T_i^U, \delta_i \mid b_{ki}, \theta\big)p(b_{ki} \mid \theta)p(\theta),$$

where θ represents the full parameter vector, which includes the fixed effects from both the sub-models, the association parameters from the survival sub-model, and the variance components from the GLMM.

Furthermore, $p\big(y_{kij} \mid b_{ki}, \theta\big)$ represents the conditional distribution for the longitudinal part of the model, which is expressed as follows:

$$p(y_{kij} \mid b_{ki}, \theta) = \exp\left\{\frac{\left[y_{kij}\psi_{kij}(b_{ki}) - c_k\{\psi_{kij}(b_{ki})\}\right]}{a_k(\varphi)} - d_k(y_{kij}\varphi)\right\},$$

where ψ_{kij} (b_{ki}) and φ denote the natural and dispersion parameters in the exponential family respectively, and $c_k(\cdot)$, $a_k(\cdot)$ and $d_k(\cdot)$ are known functions that specify the member of the exponential family.

Finally, $p(T_i, T_i^U, \delta_i \mid b_{ki})$ represents the density distribution for the survival part of the model, which is represented as follows:

$$\begin{aligned} p(T_i, T_i^U, \delta_i \mid b_{ki}, \theta) = &\{h_i(T_i \mid \boldsymbol{\mathcal{H}}_{\boldsymbol{i}}(T_i), w_i(T_i))\}^{I(\delta_i=1)} \exp\left\{-\int_0^{T_t} h_i(s \mid \boldsymbol{\mathcal{H}}_{\boldsymbol{i}}(s), w_i(s))ds\right\} \\ &\times \left\{1 - \exp\left\{-\int_0^{T_t} h_i(s \mid \boldsymbol{\mathcal{H}}_{\boldsymbol{i}}(s), w_i(s))ds\right\}\right\}^{I(\delta_i=2)} \\ &\times \left\{\exp\left\{-\int_0^{T_t} h_i(s \mid \boldsymbol{\mathcal{H}}_{\boldsymbol{i}}(s), w_i(s))ds\right\}\right. \\ &\left. - \exp\left\{-\int_0^{T_t} h_i(s \mid \boldsymbol{\mathcal{H}}_{\boldsymbol{i}}(s), w_i(s))ds\right\}\right\}^{I(\delta_i=3)}, \end{aligned}$$

where $I(\cdot)$ represents the indicator function, the parameters T_i^U , $\delta_i = 2$, and $\delta_i = 3$ are used when the data are interval censored.

CHAPTER 6:
MODEL-FITTING RESULTS

6.1 INTRODUCTION

This chapter discusses the estimation results from the standard methods addressed in Chapters 4 and 5 and their application using the IMPI trial data. Each model was also evaluated for overall fit. The study's conclusions and summary and suggestions for further research are also presented.

6.2 COX MODELLING OF THE IMPI TRIAL DATASET

This section examines the effect of some of the selected covariates on the composite, death, and constriction events using the measure of effect, namely the HR. The relationship between each covariate and the event of interest is defined using regression coefficients. Assuming all other covariates remain constant; they represent the change in the expected log of the HR for a unit change in a covariate associated with the comparison of individuals who experienced the event to those who did not experience the event.

6.2.1 Cox modelling of time-to-composite event

The standard Cox PH model was fitted and all covariates and time-dependent variables were included, which were the ECG characteristics. It is important to note that the binary time-dependent variables were only added at baseline. The results with the corresponding confidence intervals are presented in Table 6.1. The confidence intervals of all the ECG variables show an insignificant contribution to the HR in the composite event. Furthermore, they suggest that there is a 3% increased risk of composite events in patients who have a normal ECGrate compared to patients with an abnormal ECGrate. There is a 64% increased risk of a composite event for patients who have a normal ECGpri compared to abnormal patients, and there is a further 40% increased risk of a composite event for patients with a normal ECGqtc compared to patients with an abnormal ECGqrsd. Lastly, there is a reduced risk of 25% in the composite event for patients with a normal ECGqrsd compared to patients with an abnormal ECGqrsd. There is an approximately 25% reduced risk of prednisolone, which is the active arm, as compared to the placebo group. There is clearly a marked difference between the ECG characteristics and prednisolone, possibly due to differences in ECG measurements taken or the effect of confounding variables that were fitted in the Cox PH model, or even the treatment effect.

Table 6.1: Cox model for time-to-composite event

Variable	Time to composite HR (95% CI*)	p-value	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value
ECGrate	1.029 (0.682 - 1.553)	0.891	-	-	-	-	-	-
ECGpri	-	-	1.636 (0.829-3.229)	0.156	-	-	-	-
ECGqtc	-	-	-	-	1.396 (0.984 - 1.980)	0.062	-	-
ECGqrsd	-	-	-	-	-	-	0.751 (0.274 - 2.058)	0.577
Prednisolone	0.765 (0.574 - 1.019)	0.067	0.756 (0.567 - 1.006)	0.055	0.747 (0.560 - 1.123)	0.195	0.754 (0.567 - 1.005)	0.054
Pericardiocentesis at randomisation (yes)	0.749 (0.514 - 1.094)	0.135	0.798 (0.546 - 1.166)	0.243	0.819 (0.559 - 1.199)	0.305	0.792 (0.542 - 1.156)	0.227
Duration of symptoms (days)	1.004 (1.001 - 1.008)	0.022	1.005 (1.001 - 1.009)	0.016	1.005 (1.000 - 1.009)	0.015	1.005 (1.001 - 1.009)	0.016
Age (years)	1.008 (0.996 - 1.020)	0.200	1.005 (0.993 - 1.017)	0.394	1.004 (0.992 - 1.017)	0.470	1.005 (0.993 - 1.017)	0.400
NYHA class at study entry:								
II	1.405 (0.880 - 2.241)	0.153	1.463 (0.916 - 2.337)	0.111	1.495 (0.937 - 2.385)	0.091	1.462 (0.916 - 2.335)	0.111
III	1.889 (1.115 - 3.204)	0.018	1.877 (1.106 - 3.186)	0.019	1.923 (1.133 - 3.264)	0.015	1.880 (1.106 - 3.197)	0.019
IV	2.115 (1.126 - 3.972)	0.020	1.959 (1.040 - 3.693)	0.037	1.980 (1.052 - 3.727)	0.034	1.966 (1.040 - 3.711)	0.037
Creatinine (≥105)	1.141 (0.594 - 2.192)	0.692	1.231 (0.645 - 2.349)	0.528	1.199 (0.626 - 2.296)	0.582	1.230 (0.644 - 2.352)	0.530
HIV status (HIV+)	0.664 (0.469 - 0.938)	0.020	0.636 (0.449 - 0.900)	0.011	0.644 (0.455 - 0.911)	0.013	0.639 (0.453 - 0.905)	0.011
Weight (kg)	0.946 (0.831 - 1.078)	0.409	0.944 (0.829 - 1.075)	0.388	0.947 (0.831 - 1.078)	0.407	0.949 (0.834 - 1.082)	0.437
Palpable pulsus paradoxus (yes)	0.832 (0.566 - 1.222)	0.348	0.801 (0.546 - 1.177)	0.259	0.800 (0.544 - 1.177)	0.258	0.806 (0.547 - 1.188)	0.276
Systolic blood pressure (≥90)	0.496 (0.312 - 0.790)	0.003	0.486 (0.305 - 0.775)	0.002	0.495 (0.310 - 0.791)	0.003	0.489 (0.307 - 0.780)	0.003
Heart rate (≥100)	0.122 (0.016 - 0.947)	0.044	0.102 (0.013 - 0.791)	0.029	0.095 (0.012 - 0.743)	0.024	0.105 (0.013 - 0.817)	0.031
Effusion size:								
Medium (1-2 cm)	1.496 (0.582 - 3.849)	0.403	1.423 (0.549 - 3.683)	0.468	1.479 (0.568 - 3.852)	0.423	1.428 (0.551 - 3.695)	0.463
Small (<1 cm)	0.619 (0.204 - 1.882)	0.398	0.590 (0.193 - 1.808)	0.356	0.611 (0.199 - 1.877)	0.390	0.616 (0.201 - 1.885)	0.396
Chest X-ray pulmonary infiltrate (yes)	1.055 (0.621 - 1.790)	0.844	1.041 (0.614 - 1.767)	0.881	1.052 (0.618 - 1.789)	0.851	1.054 (0.621 - 1.788)	0.845
Definite TB pericarditis status (yes)	1.559 (1.067 - 2.279)	0.022	1.636 (1.117 - 2.396)	0.011	1.603 (1.093 - 2.352)	0.015	1.647 (1.124 - 2.412)	0.010
Haemoglobin (≥10)	0.617 (0.082 - 4.643)	0.638	0.568 (0.075 - 4.299)	0.584	0.606 (0.079 - 4.638)	0.630	0.584 (0.077 - 4.424)	0.603
Peripheral oedema (yes)	1.472 (1.080 - 2.007)	0.014	1.563 (1.145 - 2.133)	0.005	1.529 (1.118 - 2.092)	0.008	1.562 (1.144 - 2.132)	0.005

* CI = confidence interval

6.2.2 Cox modelling of time-to-death event

The standard Cox PH model was fitted and all covariates and binary time-dependent variables were included, which were ECGrate, ECGpri, ECGqtc, and ECGqrsd. It is worth noting that the binary time-dependent variables were only added at baseline. The results with the corresponding confidence intervals are presented in Table 6.2. The confidence intervals of all the ECG variables show an insignificant contribution to the HR in the death event. The results show that there is a 42% reduced risk of death for patients who have a normal ECGqrsd compared to patients with an abnormal ECGqrsd. There is also a 2% reduced risk of death for normal ECGrate patients compared to abnormal patients. However, there is a 19% increased risk of death for normal ECGpri patients compared to abnormal patients, and, finally, a 40% increased risk of death for normal ECGqtc patients compared to abnormal ECGqtc patients. There is an approximately 14% reduced risk of death for patients on prednisolone compared to those in the placebo group. The marked difference between the ECG characteristics and prednisolone could be because of the difference in ECG measurements or the effect of confounding variables that were fitted in the model, or even the treatment effect.

Table 6.2: Cox model for time-to-death event

Variable	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value
ECGrate	0.985 (0.596 - 1.628)	0.953	-	-	-	-	-	-
ECGpri	-	-	1.193 (0.434 - 3.276)	0.732	-	-	-	-
ECGqtc	-	-	-	-	1.396 (0.984 - 1.980)	0.062	-	-
ECGqrsd	-	-	-	-	-	-	0.579 (0.078 - 4.133)	0.449
Prednisolone	0.865 (0.687 - 1.468)	0.982	0.866 (0.567 - 1.322)	0.504	0.865 (0.566 - 1.123)	0.500	0.873 (0.571 - 1.335)	0.532
Pericardiocentesis at randomisation (yes)	0.608 (0.374 - 0.989)	0.045	0.683 (0.421 - 1.108)	0.122	0.819 (0.559 - 1.199)	0.305	0.677 (0.418 - 1.098)	0.114
Duration of symptoms (days)	0.999 (0.994 - 1.006)	0.874	0.999 (0.994 - 1.006)	0.896	1.005 (1.001 - 1.009)	0.015	0.999 (0.994 - 1.006)	0.885
Age (years)	1.012 (0.995 - 1.029)	0.172	1.009 (0.993 - 1.028)	0.261	1.004 (0.992 - 1.017)	0.470	1.009 (0.993 - 1.027)	0.267
NYHA class at study entry:								
II	1.350 (0.726 - 2.510)	0.342	1.402 (0.754 - 2.607)	0.285	1.495 (0.937 - 2.385)	0.091	1.382 (0.744 - 2.570)	0.306
III	1.861 (0.937 - 3.693)	0.076	1.749 (0.879 - 3.477)	0.111	1.923 (1.133 - 3.264)	0.015	1.734 (0.873 - 3.449)	0.116
IV	2.683 (1.259 - 5.717)	0.011	2.571 (1.204 - 5.489)	0.015	1.980 (1.052 - 3.727)	0.034	2.554 (1.193 - 5.466)	0.016
Creatinine (≥105)	2.351 (1.389 - 3.977)	0.001	2.396 (1.417 - 4.050)	0.001	1.199 (0.627 - 2.296)	0.582	2.381 (1.408 - 4.027)	0.001
HIV status (HIV+)	0.882 (0.540 - 1.441)	0.617	0.888 (0.541 - 1.457)	0.637	0.644 (0.455 - 0.912)	0.013	0.886 (0.540 - 1.453)	0.631
Weight (kg)	0.845 (0.707 - 1.010)	0.064	0.865 (0.724 - 1.034)	0.111	0.947 (0.831 - 1.078)	0.407	0.867 (0.726 - 1.036)	0.116
Palpable pulsus paradoxus (yes)	0.739 (0.453 - 1.207)	0.228	0.688 (0.421 - 1.124)	0.135	0.800 (0.544 - 1.178)	0.258	0.689 (0.421 - 1.129)	0.140
Systolic blood pressure (≥90)	0.306 (0.175 - 0.534)	0.021	0.289 (0.166 - 0.506)	0.009	0.495 (0.310 - 0.791)	0.003	0.294 (0.169 - 0.514)	0.011
Heart rate (≥100)	0.757 (0.509 - 1.128)	0.171	0.847 (0.569 - 1.260)	0.412	0.095 (0.012 - 0.743)	0.025	0.853 (0.573 - 1.269)	0.432
Effusion size:								
Medium (1-2 cm)	0.975 (0.604 - 1.574)	0.918	0.941 (0.583 - 1.521)	0.805	1.479 (0.568 - 3.852)	0.423	0.941 (0.583 - 1.519)	0.804
Small (<1 cm)	0.469 (0.192 - 1.144)	0.096	0.442 (0.180 - 1.083)	0.074	0.612 (0.199 - 1.878)	0.390	0.447 (0.182 - 1.094)	0.078
Chest X-ray pulmonary infiltrate (yes)	0.909 (0.607 - 1.363)	0.645	0.933 (0.622 - 1.399)	0.737	1.052 (0.619 - 1.789)	0.851	0.934 (0.623 - 1.399)	0.739
Definite TB pericarditis status (yes)	1.199 (0.716 - 2.007)	0.492	1.212 (0.722 - 2.036)	0.467	1.603 (1.093 - 2.352)	0.015	1.219 (0.727 - 2.048)	0.452
Haemoglobin (≥10)	0.856 (0.567 - 1.293)	0.461	0.857 (0.567 - 1.297)	0.467	0.607 (0.079 - 4.638)	0.630	0.862 (0.569 - 1.307)	0.485
Peripheral oedema (yes)	1.480 (0.989 - 2.215)	0.056	1.557 (1.039 - 2.334)	0.032	1.529 (1.118 - 2.092)	0.008	1.552 (1.036 - 2.325)	0.033

6.2.3 Cox modelling of time-to-constriction event

The standard Cox PH model was fitted and all covariates and binary time-dependent variables were included, namely ECGrate, ECGpri, ECGqtc, and ECGqrsd. It is worth noting that the binary time-dependent variables were only added at baseline. The results with the corresponding confidence intervals are presented in Table 6.3. The confidence intervals of all the ECG variables show an insignificant contribution to the HR in the constriction event.

The results show an increased risk of 2.017-fold in the constriction event for patients with a normal ECGrate compared to those with an abnormal ECGrate, and an increased risk of 1.341-fold in constriction event for patients with a normal ECGqrsd compared to patients with an abnormal ECGqrsd. However, there is a reduced risk of 31% in the constriction event for patients with a normal ECGpri compared to patients with an abnormal ECGpri and a reduced risk of 5% in the constriction event for patients with a normal ECGqtc compared to patients with an abnormal ECGqtc.

There is an approximately 59% significantly reduced risk of constriction for patients on prednisolone compared to the placebo group. It is notable that there was a marked difference between the ECG characteristics and prednisolone, which could be as a result of the difference in ECG measurements taken or the effect of confounding variables that were fitted in the Cox model, or even the treatment effect.

Table 6.3: Cox model for time-to-constriction event

Variable	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value
ECGrate	2.017 (0.844 - 4.820)	0.114	-	-	-	-	-	-
ECGpri	-	-	0.689 (0.093 - 5.101)	0.715	-	-	-	-
ECGqtc	-	-	-	-	0.953 (0.482 - 2.321)	0.889	-	-
ECGqrsd	-	-	-	-	-	-	1.341 (0.174 - 10.342)	0.778
Prednisolone	0.411 (0.210 - 0.803)	0.009	0.411 (0.211 - 0.799)	0.009	0.412 (0.211 - 0.801)	0.008	0.410 (0.211 - 0.798)	0.009
Pericardiocentesis at randomisation (yes)	0.458 (0.209 - 1.003)	0.051	0.486 (0.218 - 1.083)	0.078	0.485 (0.218 - 1.079)	0.076	0.482 (0.217 - 1.074)	0.074
Duration of symptoms (days)	0.457 (0.209 - 1.003)	0.064	1.008 (0.218 - 1.083)	0.045	1.008 (1.000 - 1.016)	0.044	1.008 (1.000 - 1.016)	0.044
Age (years)	1.000 (0.976 - 1.025)	0.984	1.001 (0.977 - 1.025)	0.956	1.000 (0.977 - 1.025)	0.975	1.000 (0.977 - 1.025)	0.964
NYHA class at study entry:								
II	1.437 (0.503 - 4.107)	0.498	1.381 (0.486 - 3.922)	0.544	1.385 (0.488 - 3.937)	0.541	1.402 (0.490 - 4.011)	0.528
III	3.511 (1.133 - 10.885)	0.029	3.176 (1.034 - 9.752)	0.044	3.157 (1.028 - 9.697)	0.045	3.203 (1.035 - 9.913)	0.043
IV	1.885 (0.395 - 8.998)	0.426	1.480 (0.316 - 6.931)	0.619	1.454 (0.310 - 6.814)	0.634	1.464 (0.313 - 6.842)	0.628
Creatinine (≥105)	1.607 (0.679 - 3.799)	0.280	1.691 (0.722 - 3.959)	0.226	1.688 (0.717 - 3.971)	0.230	1.700 (0.727 - 3.978)	0.220
HIV status (HIV+)	0.293 (0.142 - 0.603)	<0.001	0.303 (0.147 - 0.624)	<0.001	0.301 (0.146 - 0.620)	<0.001	0.299 (0.145 - 0.618)	<0.001
Weight (kg)	1.066 (0.829 - 1.371)	0.616	1.077 (0.835 - 1.389)	0.567	1.073 (0.833 - 1.383)	0.585	1.075 (0.835 - 1.386)	0.577
Palpable pulsus paradoxus (yes)	0.865 (0.397 - 1.886)	0.716	0.852 (0.387 - 1.873)	0.690	0.863 (0.394 - 1.889)	0.712	0.859 (0.393 - 1.879)	0.704
Systolic blood pressure (≥90)	2.292 (0.304 - 17.289)	0.421	2.301 (0.304 - 17.393)	0.419	2.309 (0.306 - 17.444)	0.417	2.293 (0.303 - 17.345)	0.421
Heart rate (≥100)	0.988 (0.509 -1.917)	0.972	1.349 (0.708 -2.572)	0.362	1.360 (0.712 -2.597)	0.351	1.348 (0.708 -2.569)	0.363
Effusion size:								
Medium (1-2 cm)	1.101 (0.494 - 2.454)	0.815	1.094 (0.487 - 2.459)	0.828	1.082 (0.483 - 2.425)	0.848	1.074 (0.478 - 2.414)	0.863
Small (<1 cm)	0.437 (0.095 - 1.997)	0.285	0.431 (0.094 - 1.981)	0.279	0.422 (0.092 - 1.939)	0.267	0.418 (0.091 - 1.924)	0.263
Chest X-ray pulmonary infiltrate (yes)	1.392 (0.749 - 2.588)	0.296	1.446 (0.780 - 2.681)	0.241	1.440 (0.777 - 2.672)	0.247	1.451 (0.782 - 2.693)	0.238
Definite TB pericarditis status (yes)	2.991 (1.360 - 6.578)	0.006	3.352 (1.527 - 7.359)	0.003	3.355 (1.527 - 7.369)	0.003	3.376 (1.536 - 7.419)	0.002
Haemoglobin (≥10)	1.583 (0.794 - 3.157)	0.192	1.576 (0.784 - 3.171)	0.202	1.581 (0.785 - 3.182)	0.199	1.565 (0.775 - 3.159)	0.211
Peripheral oedema (yes)	1.348 (0.693 - 2.623)	0.379	1.491 (0.764 - 2.909)	0.241	1.484 (0.758 - 2.904)	0.249	1.486 (0.761 - 2.902)	0.246

6.2.4 Cox modelling overall model fit: Cox-Snell residual plots

This subsection shows the Cox-Snell residual plots that were used to assess the overall fit of the Cox model. These plots provide insight into how well a Cox model fits. By examining them, it can be determined whether the Cox model is a good fit.

6.2.4.1 Composite event Cox-Snell residual plots for the standard Cox model

As shown in the figures below, residuals were plotted on the x-axis, with the integrated hazards of the residuals plotted on the y-axis. Ideally, the curve of the Cox-Snell residuals against the integrated hazard should fall roughly along the 45-degree line if the model works well. Figures 6.1 to 6.4 show Cox-Snell residual plots for four binary ECG characteristics, which correspond to the Cox PH model considering the composite event outcome. The figures show that the line plots deviate too much from the reference line for all four binary ECG characteristics; the Cox PH models therefore did not fit the data well.

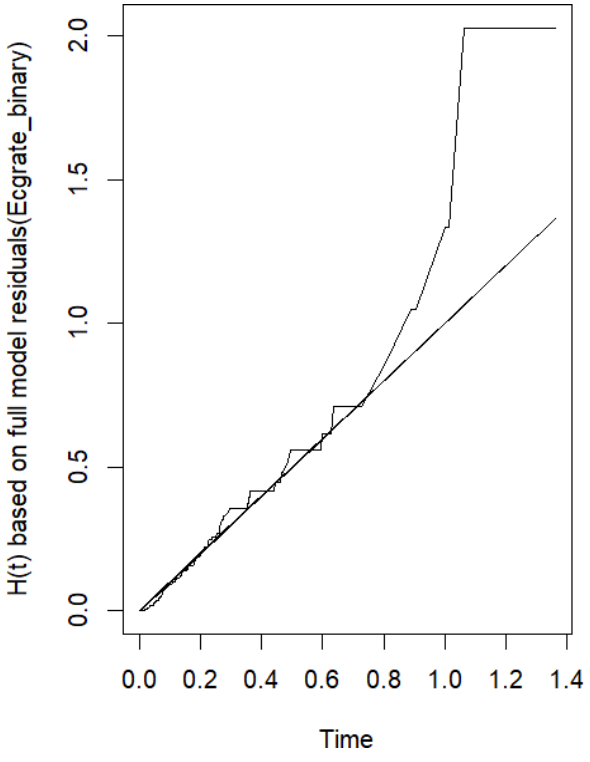

Figure 6.1: Residual plot corresponds to Cox model ECGrate

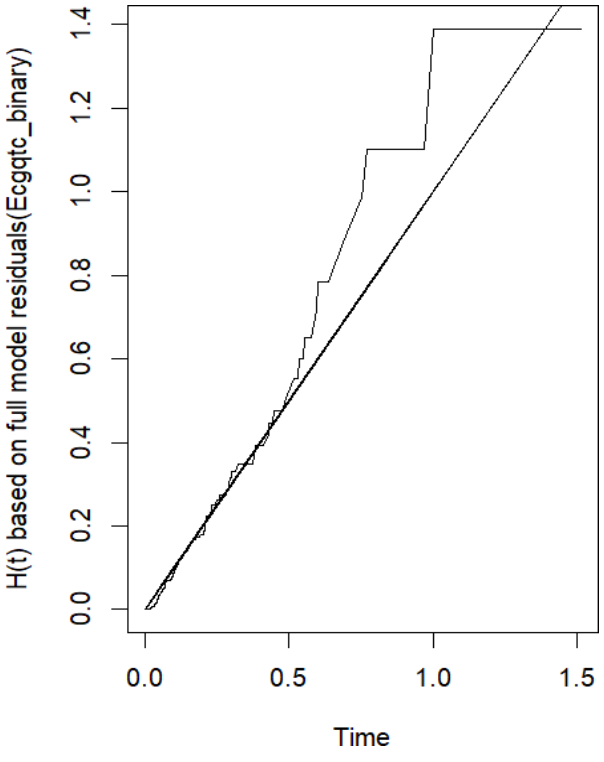

Figure 6.2: Residual plot corresponds to Cox model ECGqtc

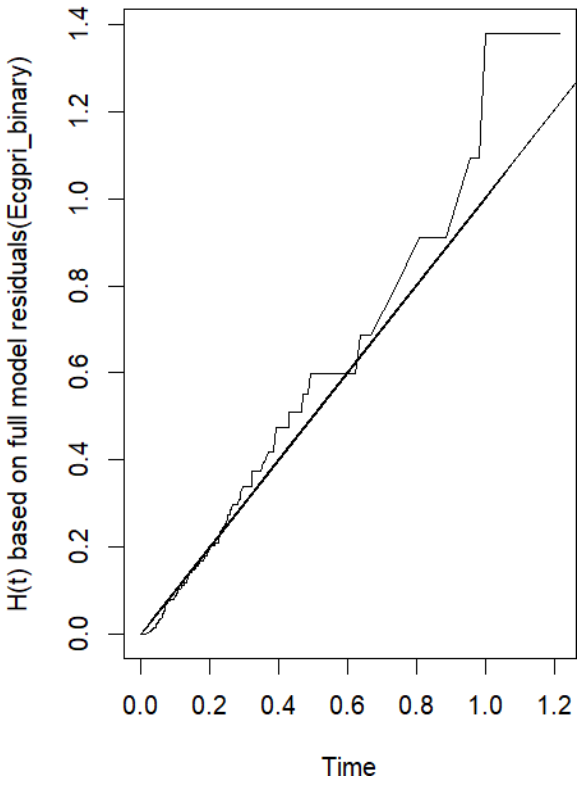

Figure 6.3: Residual plot corresponds to Cox model ECGpri

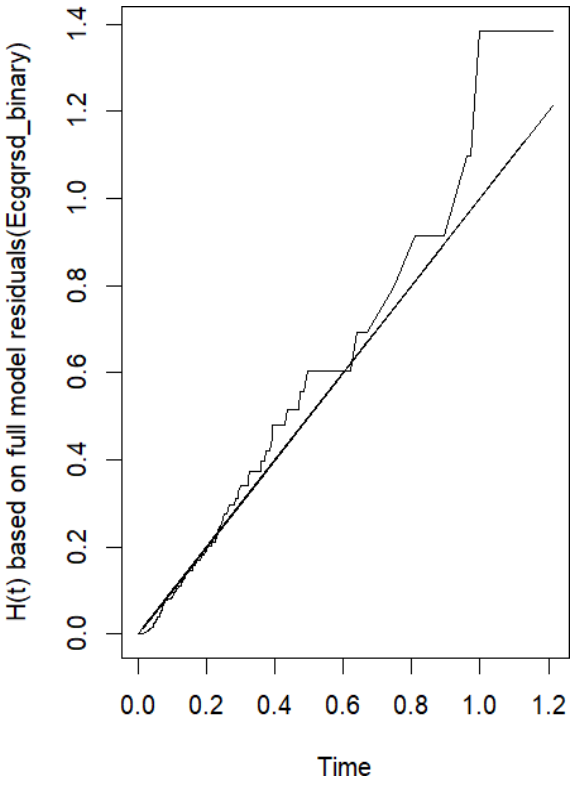

Figure 6.4: Residual plot corresponds to Cox model ECGqrsd

6.2.4.2 Death event Cox-Snell residual plots for standard Cox model

Figures 6.5 to 6.8 show the Cox-Snell residual plots of four binary ECG characteristics, which correspond to the Cox PH model considering the death event outcome. The figures show that the line plots deviate too much from the reference line for all four binary ECG characteristics, which means that the Cox PH models did not fit the data well.

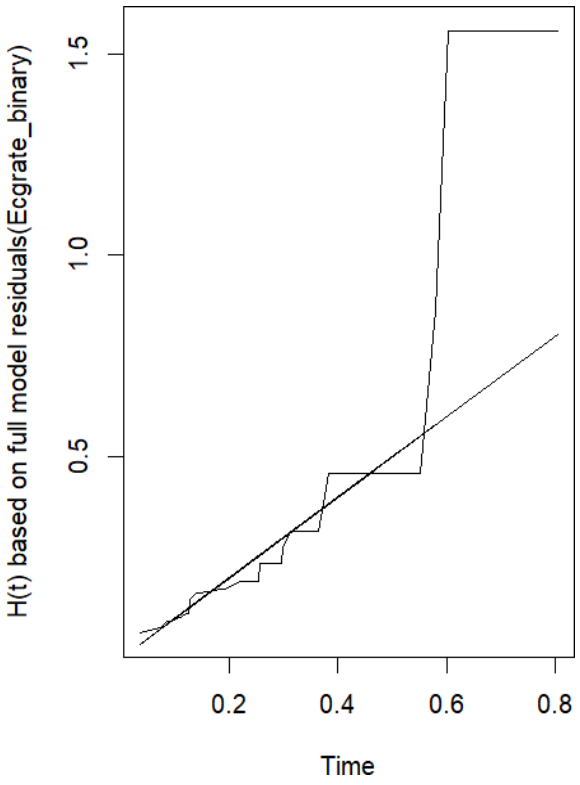

Figure 6.5: Residual plot corresponds to Cox model ECGrate

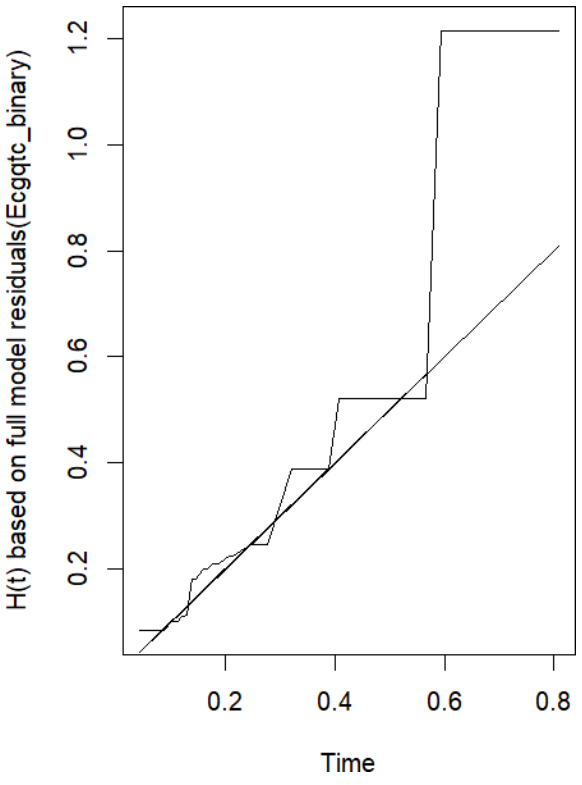

Figure 6.6: Residual plot corresponds to Cox model ECGqtc

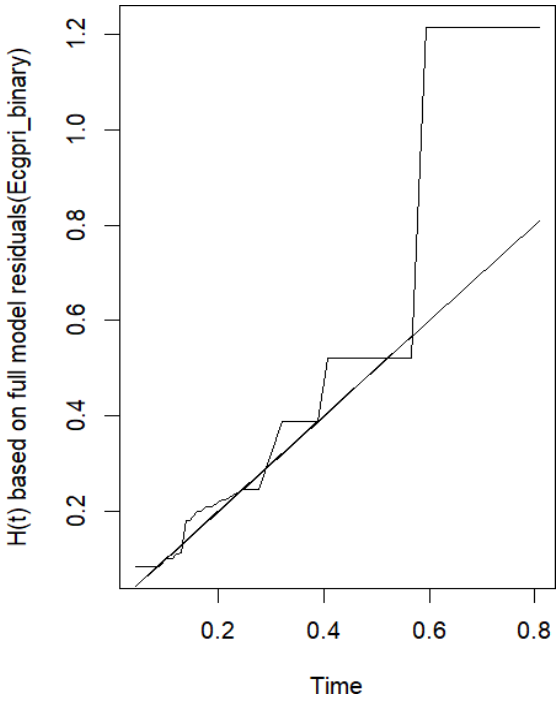

Figure 6.7: Residual plot corresponds to Cox model ECGpri

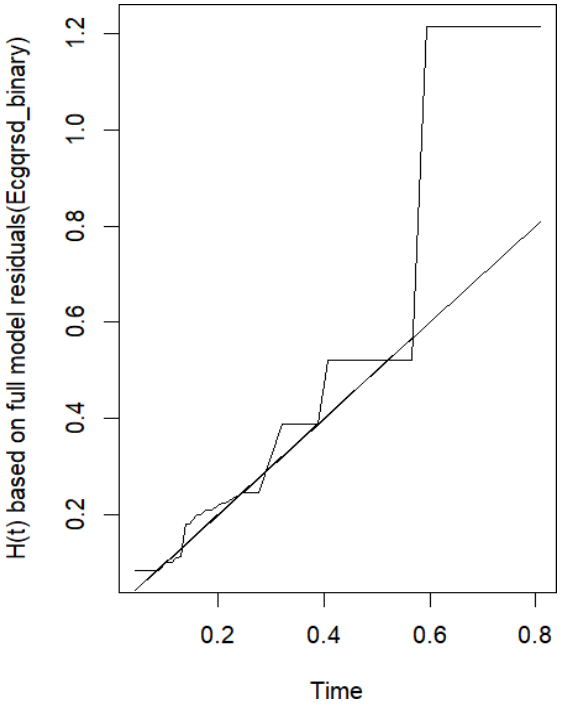

Figure 6.8: Residual plot corresponds to Cox model ECGqrsd

6.2.4.3 Constriction event Cox-Snell residual plots for the Cox PH model

Figures 6.9 to Figure 6.12 show the Cox-Snell residual plot corresponding to the Cox PH model of the ECG characteristics considering the constriction event outcome. The figures also show that the line plots deviate too much from the reference line for all four binary ECG characteristics, which means that the Cox PH models did not fit the data well.

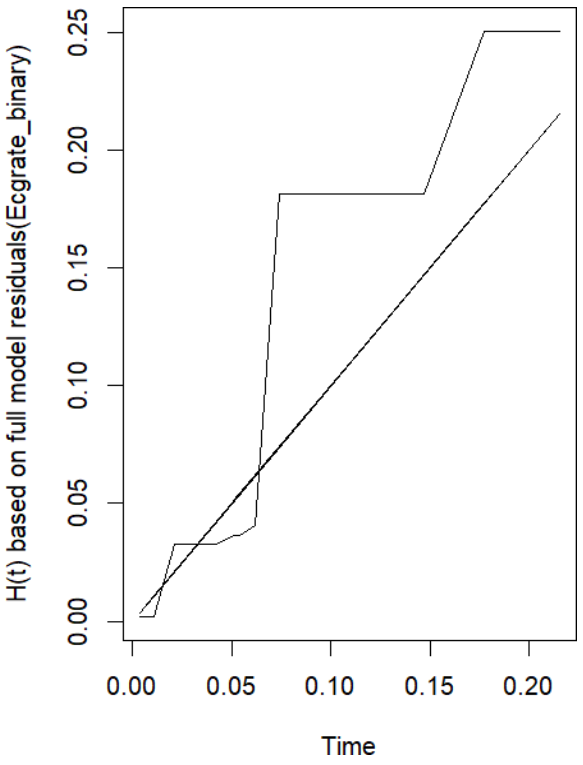

Figure 6.9: Residual plot corresponds to Cox model ECGrate

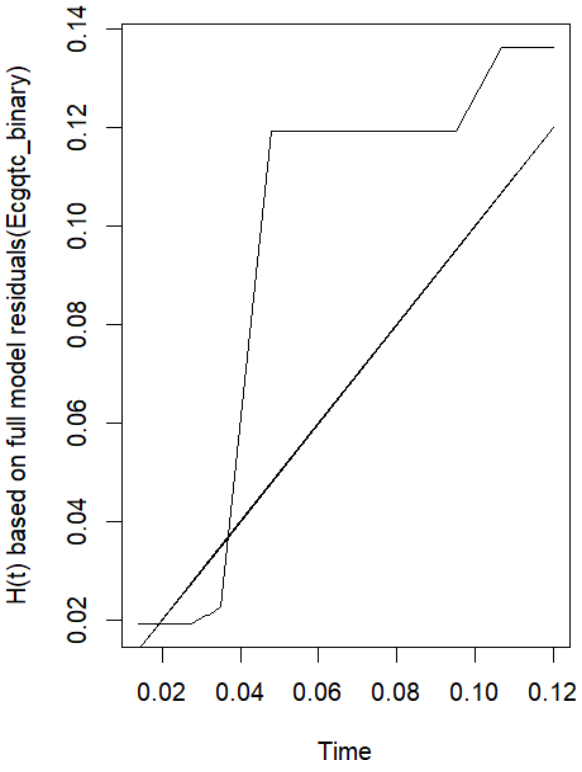

Figure 6.10: Residual plot corresponds to Cox model ECGqtc

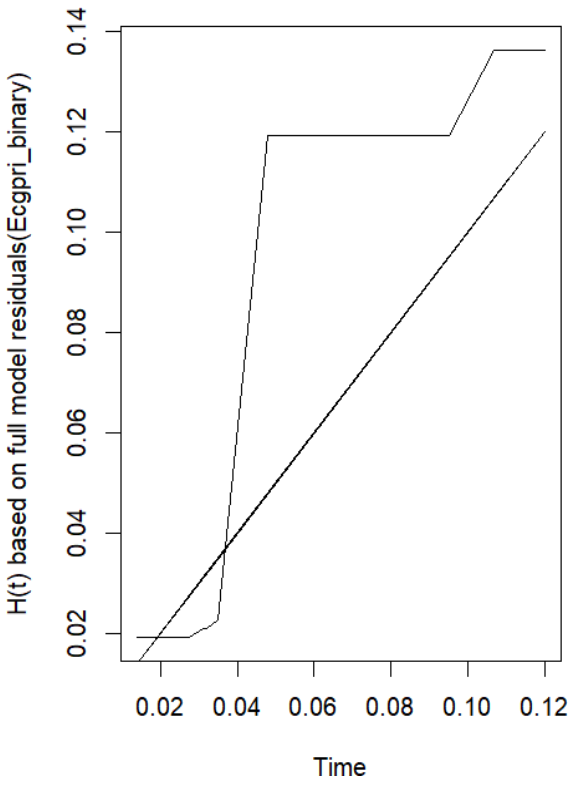

Figure 6.11: Residual plot corresponds to Cox model ECGpri

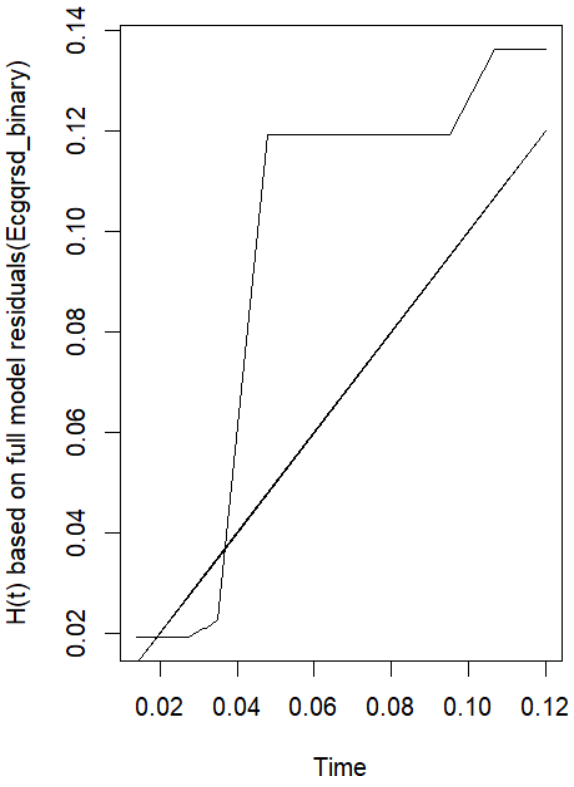

Figure 6.12: Residual plot corresponds to Cox model ECGqrsd

6.3 EXTENDED COX MODELLING OF IMPI TRIAL DATASET

As explained in Section 4.4, the extended Cox model is an extension of the Cox PH regression model that fits exogenous time-varying covariates. Therefore, when fitting this model, time-updated data were used, whereas only baseline observations were used in the standard Cox PH model. Covariates at baseline and time-varying covariates are part of the linear predictor ζ_i.

For the IMPI trial dataset, an extended Cox analysis selected baseline covariates specified in Chapter 2 for analysis in time-to-composite, time-to-death, and time-to-constriction events, respectively. Moreover, the four ECG variables are the binary time-varying covariates in the model.

This section examines the effect of some of the selected covariates on the composite, death, and constriction events using the measure of effect, namely the HR. The relationship between each covariate and the event of interest is defined using regression coefficients. Assuming all other covariates remain constant, they represent the change in the expected log of the HR for a unit change in a covariate associated with the comparison of individuals who experienced the

event to those who did not experience the event. For inferring the clinical significance of a covariate, determining effect sizes helps to know how much actual change a covariate makes to a response variable (Hazra, 2017).

For each time-to-event analysis, 95% confidence intervals were calculated. The results below refer to the extended Cox models that were fitted, including each of the binary time-varying covariates (i.e., ECGrate, ECGpri, ECGqtc, and ECGqrsd), one at a time to allow for comparison with future joint models. Tables 6.1 to 6.3 show the results of the three extended Cox models that examined the influence of baseline and time-varying covariates on the hazard rate for the composite outcome, death, and constriction, respectively.

6.3.1 Extended Cox modelling of time-to-composite event

An extended Cox model was fitted and all covariates and time-dependent variables were included. The results are shown in Table 6.4, with their corresponding confidence intervals. The confidence intervals for all the covariates, except for duration of symptoms (days), definite TB pericarditis status (yes), systolic blood pressure (≥ 90), heart rate (≥ 100), and peripheral oedema (yes), showed a significant contribution to the HR in the composite event.

Statistical significance was observed for two NYHA class categories at study entry but not for one class on the composite event. It was only significant for two categories of the time-to-composite event but not one. Therefore, the entire covariate was taken into account as it was not sufficient to exclude only the categories with an insignificant difference from the analysis. In addition, excluding a category would combine the insignificant level with the reference level, the interpretation would change, or it would not make sense based on the nature of the covariate.

Table 6.4: Extended Cox model for time-to-composite event

Variable	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value
ECGrate	2.151 (0.239 - 19.314)	0.494	-	-	-	-	-	-
ECGpri	-	-	2.053 (0.955 - 4.415)	0.065	-	-	-	-
ECGqtc	-	-	-	-	1.394 (0.984 - 1.977)	0.062	-	-
ECGqrsd	-	-	-	-	-	-	0.724 (0.503 - 3.788)	0.531
Prednisolone	0.821 (0.615 - 1.094)	0.178	0.812 (0.611 - 1.081)	0.154	0.809 (0.608 - 1.076)	0.144	0.806 (0.604 - 1.073)	0.140
Pericardiocentesis at randomisation (yes)	0.783 (0.535 - 1.145)	0.207	0.780 (0.533 - 1.141)	0.201	0.804 (0.548 - 1.179)	0.264	0.782 (0.534 - 1.144)	0.204
Duration of symptoms (days)	1.005 (1.001 - 1.009)	0.009	1.005 (1.001 - 1.009)	0.008	1.005 (1.001 - 1.009)	0.009	1.005 (1.001 - 1.009)	0.009
Age (years)	0.996 (0.992 - 1.016)	0.503	0.997 (0.992 - 1.015)	0.584	1.003 (0.991 - 1.015)	0.671	0.997 (0.992 - 1.015)	0.563
NYHA Class at study entry:								
II	1.327 (0.835 - 2.109)	0.232	1.320 (0.830 - 2.099)	0.240	1.347 (0.848 - 2.139)	0.207	1.333 (0.838 - 2.118)	0.226
III	1.874 (1.102 - 3.186)	0.020	1.815 (1.067 - 3.086)	0.028	1.869 (1.099 - 3.178)	0.021	1.876 (1.104 - 3.188)	0.020
IV	2.576 (1.369 - 4.847)	0.003	2.521 (1.338 - 4.751)	0.004	2.543 (1.353 - 4.785)	0.004	2.559 (1.361 - 4.815)	0.004
Creatinine (≥105)	1.678 (0.870 - 3.236)	0.123	1.628 (0.847 - 3.130)	0.144	1.609 (0.834 - 3.105)	0.156	1.661 (0.861 - 3.198)	0.128
HIV status (HIV+)	0.755 (0.537 - 1.061)	0.105	0.751 (0.533 - 1.057)	0.101	0.759 (0.540 - 1.069)	0.114	0.756 (0.537 - 1.064)	0.108
Weight (kg)	1.0622 (0.830 - 1.067)	0.346	0.934 (0.824 - 1.059)	0.287	0.942 (0.831 - 1.069)	0.353	0.938 (0.827 - 1.064)	0.320
Palpable pulsus paradoxus (yes)	0.996 (0.680 - 1.485)	0.982	1.000 (0.676 - 1.476)	0.997	1.005 (0.673 - 1.472)	0.980	1.003 (0.675 - 1.475)	0.990
Systolic blood pressure (≥90)	0.597 (0.375 - 0.950)	0.029	0.591 (0.371 - 0.940)	0.026	0.593 (0.373 - 0.944)	0.027	0.593 (0.373 - 0.944)	0.028
Heart rate (≥100)	0.094 (0.012 - 0.746)	0.025	0.091 (0.011 - 0.717)	0.023	0.084 (0.011 - 0.668)	0.019	0.093 (0.012 - 0.738)	0.025
Effusion size:								
Medium (1-2 cm)	0.705 (0.548 - 3.671)	0.470	0.716 (0.536 - 3.612)	0.497	0.664 (0.578 - 3.931	0.402	0.703 (0.548 - 3.670)	0.469
Small (<1 cm)	1.567 (0.210 - 1.940)	0.429	1.613 (0.203 - 1.891)	0.401	1.557 (0.210 - 1.963)	0.437	1.575 (0.208 - 1.935)	0.424
Chest X-ray pulmonary infiltrate (yes)	0.976 (0.604 - 1.739)	0.928	0.974 (0.606 - 1.739)	0.923	0.988 (0.595 - 1.722)	0.964	0.976 (0.604 - 1.738)	0.928
Definite TB pericarditis status (yes)	1.517 (1.039 - 2.214)	0.031	1.521 (1.043 - 2.219)	0.029	1.486 (1.017 - 2.171)	0.040	1.534 (1.051 - 2.239)	0.027
Haemoglobin (≥10)	1.718 (0.077 - 4.388)	0.600	1.797 (0.074 - 4.202)	0.570	1.719 (0.077 - 4.17)	0.600	1.791 (0.074 - 4.213)	0.572
Peripheral oedema (yes)	1.529 (1.123 - 2.081)	0.006	1.539 (1.131 - 2.096)	0.006	1.498 (1.099 - 2.042)	0.011	1.529 (1.124 - 2.083)	0.007

The association trends in the extended Cox model are similar to those in the Cox PH model. However, there is a distinction while interpreting the binary time-varying covariates, such as ECGrate, ECGpri, ECGqtc, and ECGqrsd in the extended Cox model. There is an increased risk of 2.151-fold for patients with a normal ECGrate compared to patients with an abnormal rate. There is a reduced risk of 18% for patients who are on prednisolone compared to the placebo group and who have a normal ECGrate compared to an abnormal ECGrate in the composite event.

Patients with a normal ECGqrsd have a 28% reduced risk of the composite event compared to patients with an abnormal ECGqrsd rate. Furthermore, patients with a normal ECGpri have an increased risk of 2.053-fold compared to abnormal ECGpri patients. Patients with a normal ECGqtc rate have a 39% risk of experiencing either death, constriction, or tamponade compared to those in the abnormal group.

There is also a reduced risk of 19% for patients who are on prednisolone compared to the placebo group and who have a normal ECGpri, ECGqtc, and ECGqrsd compared to patients with an abnormal ECGpri, ECGqtc, and ECGqrsd in the composite event. It is important to note that there is a notable marked difference between the ECG characteristics and prednisolone, which could be due to differences in ECG measurements taken or the effect of "confounding variables" that were fitted in the model, or even the treatment effect.

Holding the other covariates constant, the expected risk of the composite event of the duration of symptoms is increased by a factor of 0.995. In addition, a patient with peripheral oedema has an increased expected risk of either death, constriction, or tamponade by a factor of 0.688 compared to those without it. There is an expected increased risk of the composite event of 10.669-fold for patients with a heart rate ≥100 compared to the ones with a heart rate <100. Similarly, for patients with systolic blood pressure (≥90), there is an expected increase in the risk of the composite event by a factor of 1.697 compared to those with systolic blood pressure (<90). The duration of symptoms in patients shows an increased expected risk of the composite event of 99.5%. On the other hand, patients in NYHA classes III and IV have an increased expected risk of the composite event by factors of 0.544 and 0.396 respectively, compared to other classes (I and II) that do not make a significant contribution to HR.

6.3.2 Extended Cox modelling of time-to-death event

The NYHA class, systolic blood pressure (≥90), and peripheral oedema show a statistically significant association with event death. The results in Table 6.5 show that at any given time *t*, there is a reduction of 33% in the risk of death for patients with a normal ECGpri compared to those with an abnormal rate. There is a 2% reduction in the hazard of death for patients with a normal ECGqrsd rate compared to those without it. Moreover, patients who have a normal ECGrate have an 11% lower risk of experiencing death compared to those with an abnormal rate. Patients who have a normal ECGqtc also have a 15% chance of experiencing death compared to the ones with an abnormal rate. There is a lowered risk of approximately 2% for prednisolone compared to the placebo group for patients who have a normal ECGrate, ECGpri, ECGqtc, and ECGqrsd compared to all abnormal ECG characteristics in the death event. There is a slight difference in ECG characteristic measurements and prednisolone, which is the active arm. The differences in ECG measurements taken, the effect of "confounding variables" that were fitted in the extended Cox model, or even the treatment effect, could be the reason.

Table 6.5: Extended Cox model for time-to-death event

Variable	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value
ECGrate	0.887 (0.737 - 1.724)	0.580	-	-	-	-	-	-
ECGpri	-	-	0.667 (0.544 - 4.128)	0.434	-	-	-	-
ECGqtc	-	-	-	-	0.853 (0.720 - 1.910)	0.522	-	-
ECGqrsd	-	-	-	-	-	-	0.980 (0.248 - 4.207)	0.977
Prednisolone	0.980 (0.701 - 1.485)	0.917	0.979 (0.703 - 1.486)	0.910	0.979 (0.702 - 1.485)	0.913	0.981 (0.700 - 1.485)	0.919
Pericardiocentesis at randomisation (yes)	0.663 (0.409 - 1.075)	0.096	0.665 (0.411 - 1.080)	0.099	0.669 (0.413 - 1.087)	0.104	0.664 (0.409 - 1.077)	0.097
Duration of symptoms (days)	0.999 (0.994 - 1.006)	0.923	1.000 (0.994 - 1.006)	0.923	1.000 (0.994 - 1.006)	0.936	1.000 (0.994 - 1.006)	0.921
Age (years)	1.011 (0.994 - 1.028)	0.198	1.011 (0.994 - 1.028)	0.204	1.011 (0.994 - 1.028)	0.208	1.011 (0.995 - 1.028)	0.192
NYHA class at study entry:								
II	0.831 (0.648 - 2.233)	0.558	0.822 (0.655 - 2.257)	0.536	0.817 (0.659 - 2.274)	0.521	0.824 (0.654 - 2.253)	0.539
III	1.746 (0.872 - 3.455)	0.117	1.721 (0.865 - 3.423)	0.122	1.733 (0.870 - 3.450)	0.118	1.729 (0.870 - 3.448)	0.118
IV	3.096 (1.442 - 6.649)	0.004	3.079 (1.434 - 6.614)	0.004	3.068 (1.425 - 6.575)	0.004	3.046 (1.433 - 6.627)	0.004
Creatinine (≥105)	2.177 (1.294 - 3.659)	0.003	2.183 (1.298 - 3.671)	0.003	2.169 (1.291 - 3.647)	0.003	2.171 (1.294 - 3.660)	0.004
HIV status (HIV+)	0.905 (0.679 - 1.802)	0.687	0.906 (0.676 - 1.802)	0.692	0.894 (0.685 - 1.828)	0.653	0.905 (0.678 - 1.803)	0.688
Weight (kg)	0.865 (0.731 - 1.020)	0.087	0.864 (0.731 - 1.021)	0.083	0.866 (0.734 - 1.024)	0.091	0.866 (0.733 - 1.022)	0.089
Palpable pulsus paradoxus (yes)	1.125 (0.542 - 1.458)	0.641	1.139 (0.536 - 1.438)	0.605	1.125 (0.542 - 1.459)	0.642	1.131 (0.539 - 1.450)	0.625
Systolic blood pressure (≥90)	0.398 (0.227 - 0.699)	0.001	0.394 (0.224 - 0.689)	0.001	0.394 (0.225 - 0.691)	0.001	0.391 (0.225 - 0.695)	0.001
Heart rate (≥100)	1.078 (0.624 - 1.380)	0.712	1.041 (0.650 - 1.420)	0.840	1.057 (0.640 - 1.397)	0.780	1.053 (0.643 - 1.402)	0.794
Effusion size:								
Medium (1-2 cm)	0.882 (0.696 - 1.850)	0.614	0.879 (0.699 - 1.854)	0.603	0.879 (0.698 - 1.855)	0.604	0.871 (0.704 - 1.874)	0.609
Small (<1 cm)	0.599 (0.244 - 1.458)	0.261	0.592 (0.242 - 1.444)	0.250	0.588 (0.240 - 1.441)	0.246	0.598 (0.248 - 1.483)	0.255
Chest X-ray pulmonary infiltrate (yes)	1.033 (0.649 - 1.446)	0.876	1.025 (0.654 - 1.457)	0.905	1.042 (0.642 - 1.437)	0.843	1.016 (0.660 - 1.470)	0.896
Definite TB pericarditis status (yes)	0.820 (0.731 - 2.035)	0.448	0.822 (0.729 - 2.032)	0.452	0.824 (0.726 - 2.028)	0.461	0.822 (0.730 - 2.030)	0.426
Haemoglobin (≥10)	1.041 (0.633 - 1.457)	0.850	1.042 (0.633 - 1.457)	0.848	1.045 (0.626 - 1.439)	0.805	1.060 (0.622 - 1.430)	0.821
Peripheral oedema (yes)	1.515 (1.012 - 2.267)	0.050	1.517 (1.016 - 2.276)	0.043	1.509 (1.008 - 2.261)	0.046	1.519 (1.015 - 2.276)	0.043

6.3.3 Extended Cox modelling of time to constriction

The results in Table 6.6 indicate that prednisolone, pericardiocentesis at randomisation, duration of symptoms, HIV status (HIV+), and definite TB pericarditis status show a significant association with a constriction event. Table 6.6 further shows an insignificant 1% reduction of ECGqtc in the risk of a constriction event at given time points *t* for patients who have a normal ECGqtc rate compared to those who have an abnormal rate. Moreover, the results in Table 6.6 show that at any given time *t*, a unit increase in ECGqrsd results in a 66.4% decrease in the risk-of-constriction event for patients who have a normal ECGqrsd rate compared to those who have an abnormal rate. ECGrate is highly associated with the risk of constriction for someone who has a normal ECGrate and has a 61.9% reduction rate compared to someone with an abnormal ECGrate. Patients who have a normal ECGpri have a 50% risk of constriction compared to those with an abnormal ECGpri. There is a significantly reduced risk of 59% for prednisolone compared to the placebo group for patients with a normal ECGrate compared to those with an abnormal ECGrate. There is also a lowered risk of 58% for prednisolone compared to the placebo group for patients who have a normal ECGpri and ECGqtc rate compared to those with an abnormal ECGpri and ECGqtc rate. Additionally, there is a reduced risk of 60% for prednisolone compared to the placebo group for patients with a normal ECGqrsd compared to abnormal ECGqrsd. There is a notable marked difference between the ECG characteristics and prednisolone, which could have resulted from the difference in ECG measurements taken or the effect of “confounding variables” that were fitted in the extended model, or even the treatment effect.

Table 6.6: Extended Cox model for time-to-constriction event

Variable	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value
ECGrate	0.409 (0.208 - 0.806)	0.0027	-	-	-	-	-	-
ECGpri	-	-	1.463 (0.090 - 5.214)	0.714	-	-	-	-
ECGqtc	-	-	-	-	0.900 (0.506 - 2.440)	0.794	-	-
ECGqrsd	-	-	-	-	-	-	0.336 (0.369 - 24.103)	0.306
Prednisolone	0.409 (0.208 - 0.806)	0.009	0.421 (0.214 - 0.825)	0.012	0.424 (0.217 - 0.831)	0.012	0.404 (0.203 - 0.803)	0.010
Pericardiocentesis at randomisation (yes)	0.397 (0.182 - 0.868)	0.021	0.433 (0.196 - 0.950)	0.037	0.427 (0.194 - 0.936)	0.034	0.416 (0.189 - 0.916)	0.030
Duration of symptoms (days)	1.008 (1.001 - 1.015)	0.030	1.008 (1.001 - 1.016)	0.022	1.008 (1.001 - 1.016)	0.022	0.991 (1.001 - 1.016)	0.019
Age (years)	0.999 (0.977 - 1.027)	0.912	0.997 (0.979 - 1.029)	0.795	0.997 (0.978 - 1.028)	0.830	0.997 (0.978 - 1.028)	0.791
NYHA class at study entry:								
II	0.681 (0.515 - 4.187)	0.472	0.681 (0.519 - 4.152)	0.469	0.673 (0.525 - 4.195)	0.456	0.679 (0.521 - 4.166)	0.466
III	3.763 (1.199 - 11.813)	0.023	3.614 (1.165 - 11.200)	0.026	3.582 (1.155 - 11.069)	0.027	0.277 (1.169 - 11.187)	0.026
IV	2.533 (0.529 - 12.122)	0.245	2.429 (0.511 - 11.311)	0.267	2.374 (0.505 - 11.241)	0.273	0.425 (0.499 - 11.107)	0.280
Creatinine (≥105)	0.659 (0.643 - 3.584)	0.341	0.664 (0.640 - 3.548)	0.348	0.669 (0.632 - 3.535)	0.360	0.668 (0.636 - 3.521)	0.356
HIV status (HIV+)	0.365 (0.178 - 0.753)	0.006	0.396 (0.192 - 0.818)	0.012	0.393 (0.190 - 0.811)	0.012	2.508 (0.193 - 0.825)	0.013
Weight (kg)	0.935 (0.833 - 1.372)	0.600	0.938 (0.829 - 1.371	0.616	0.939 (0.829 - 1.370)	0.621	0.941 (0.826 - 1.366)	0.636
Palpable pulsus paradoxus (yes)	0.992 (0.458 - 2.218)	0.985	0.944 (0.483 - 2.326)	0.886	0.928 (0.494 - 2.350)	0.851	0.41 (0.487 - 2.318)	0.879
Systolic blood pressure (≥90)	2.806 (0.371 - 21.218)	0.317	2.951 (0.390 - 22.304)	0.295	0.337 (0.392 - 22.450)	0.292	0.345 (0.384 - 21.938)	0.302
Heart rate (≥100)	0.981 (0.521 -1.995)	0.955	0.788 (0.663 -2.437)	0.470	0.779 (0.669 -2.463)	0.454	0.769 (0.680 - 2.490)	0.427
Effusion size:								
Medium (1-2 cm)	0.906 (0.489 - 2.492)	0.811	0.877 (0.502 - 2.591)	0.755	0.894 (0.496 - 2.522)	0.787	0.900 (0.492 - 2.505)	0.800
Small (<1 cm)	1.984 (0.111 - 2.299)	0.376	1.997 (0.110 - 2.277)	0.371	2.023 (0.109 - 2.254)	0.363	2.040 (0.108 - 2.228)	0.356
Chest X-ray pulmonary infiltrate (yes)	0.779 (0.690 - 2.387)	0.431	1.382 (0.743 - 2.566)	0.307	0.731 (0.735 - 2.542)	0.323	0.716 (0.750 - 2.602)	0.292
Definite TB pericarditis status (yes)	2.760 (1.254 - 6.075)	0.012	2.942 (1.342 - 6.463)	0.007	0.337 (1.351 - 6.508)	0.007	0.327 (1.386 - 6.754)	0.006
Haemoglobin (≥10)	1.755 (0.873 - 3.530)	0.115	1.716 (0.850 - 3.471)	0.132	0.582 (0.850 - 3.476)	0.132	0.578 (0.855 - 3.497)	0.127
Peripheral oedema (yes)	0.748 (0.680 - 2.624)	0.400	1.442 (0.738 - 2.832)	0.283	0.701 (0.727 - 2.803)	0.301	0.710 (0.717 - 2.767)	0.320

6.3.4 Extended Cox modelling overall model fit: Cox-Snell residual plots

This subsection shows the Cox-Snell residual plots that were used to assess the overall fit of the Cox model. These plots provide insight into how well a Cox model fits. Examining them can help to determine whether the Cox model is a good fit.

6.3.4.1 Composite event Cox-Snell residual plots for extended Cox model

The lines of the plots in Figures 6.13 to 6.16 do not deviate too much from the reference line for the extended Cox model when looking at the composite event outcome. The extended Cox model therefore provides a reasonably good fit for the data.

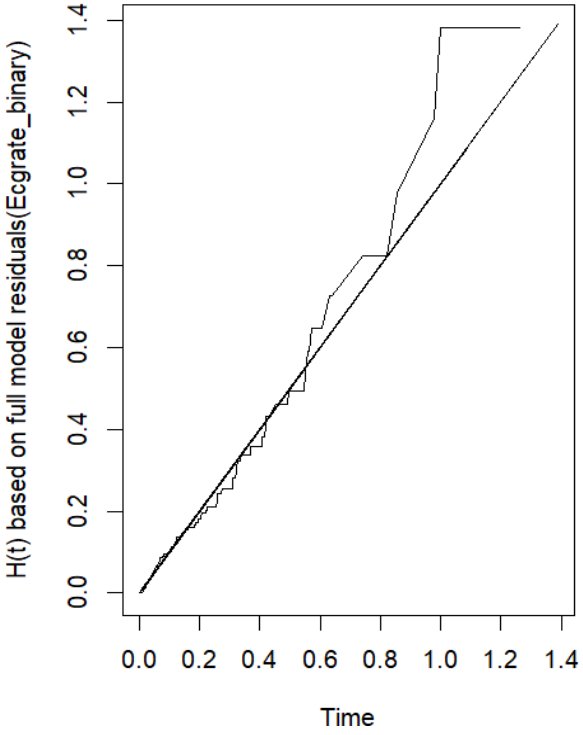

Figure 6.13: Residual plot corresponds to extended Cox model ECGrate

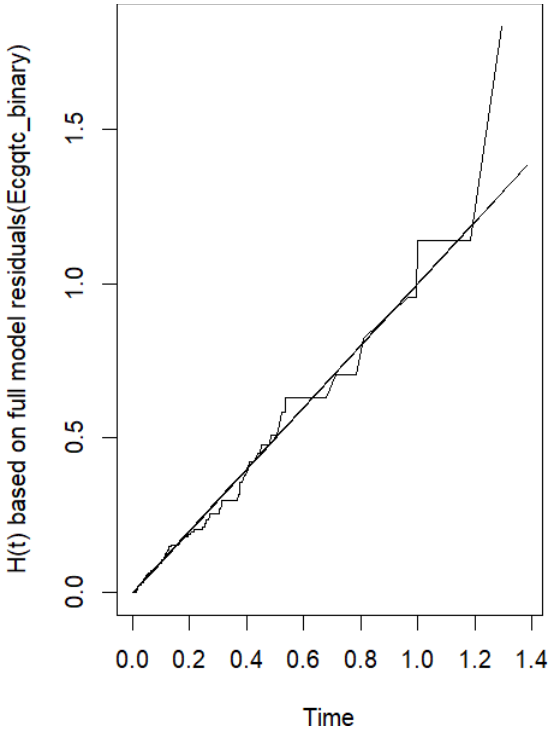

Figure 6.14: Residual plot corresponds to extended Cox model ECGqtc

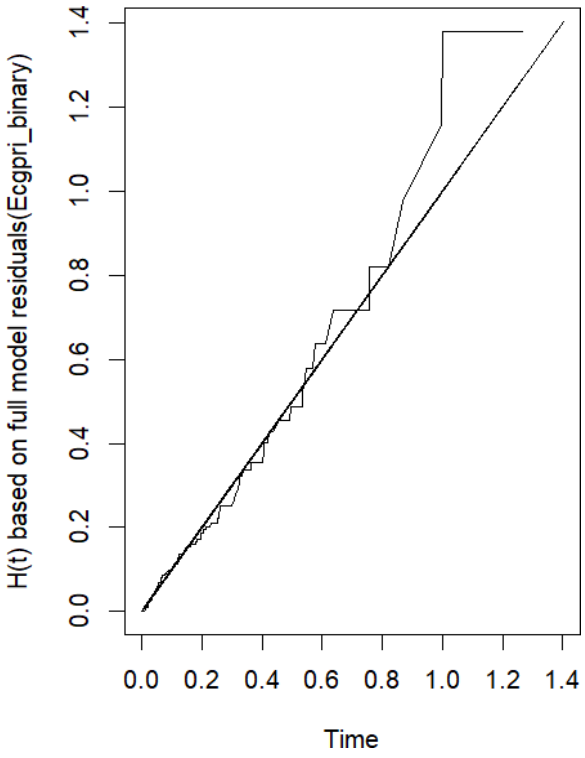

Figure 6.15: Residual plot corresponds to extended Cox model ECGpri

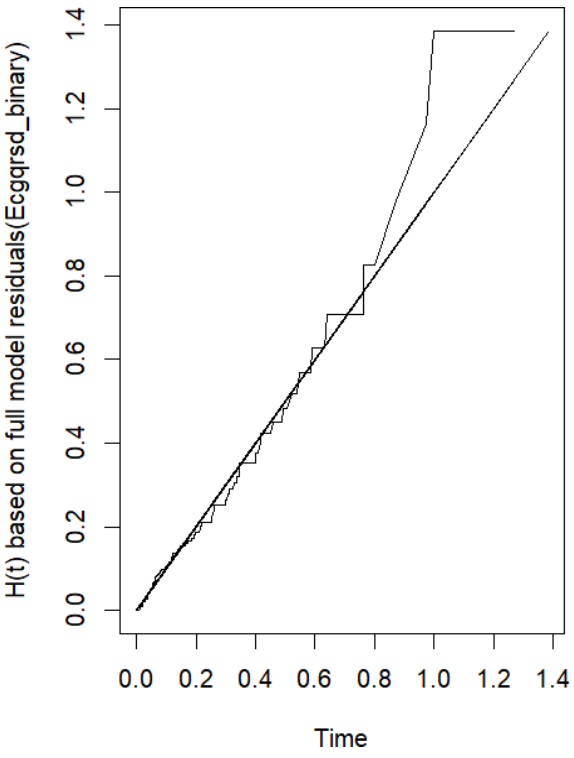

Figure 6.16: Residual plot corresponds to extended Cox model ECGqrsd

6.3.4.2 Death event Cox-Snell residual plots for extended Cox model

Figures 6.17 to 6.20 show the Cox-Snell residuals corresponding to the extended Cox model for four binary ECG characteristics when looking at the death event outcome. The line plots do not deviate too much from the reference line, which means the extended Cox model provides a reasonably good fit.

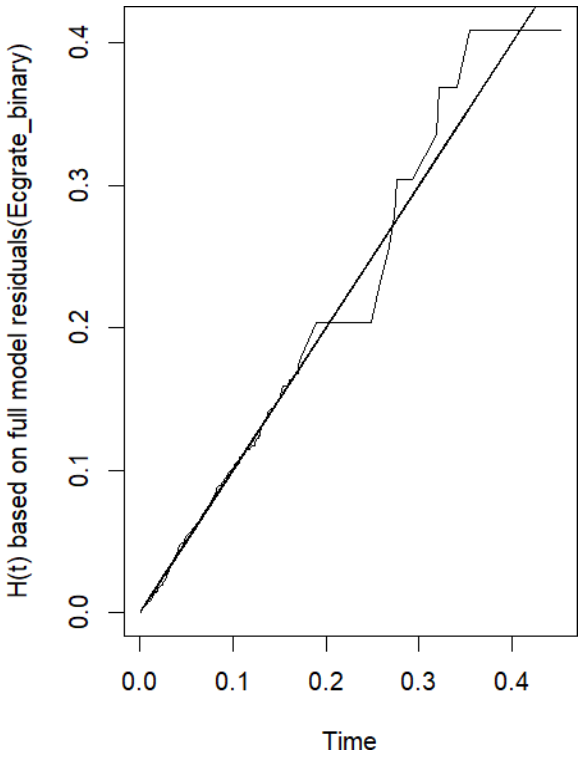

Figure 6.17: Residual plot corresponds to extended Cox model ECGrate

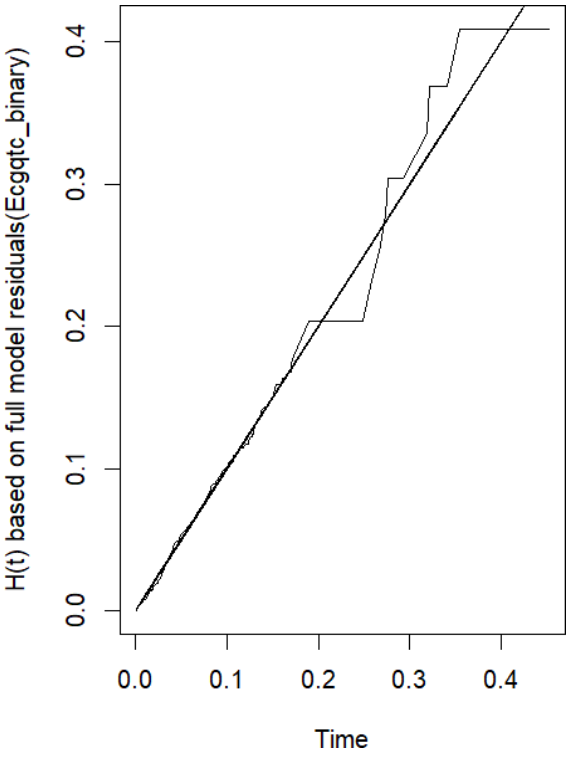

Figure 6.18: Residual plot corresponds to extended Cox model ECGqtc

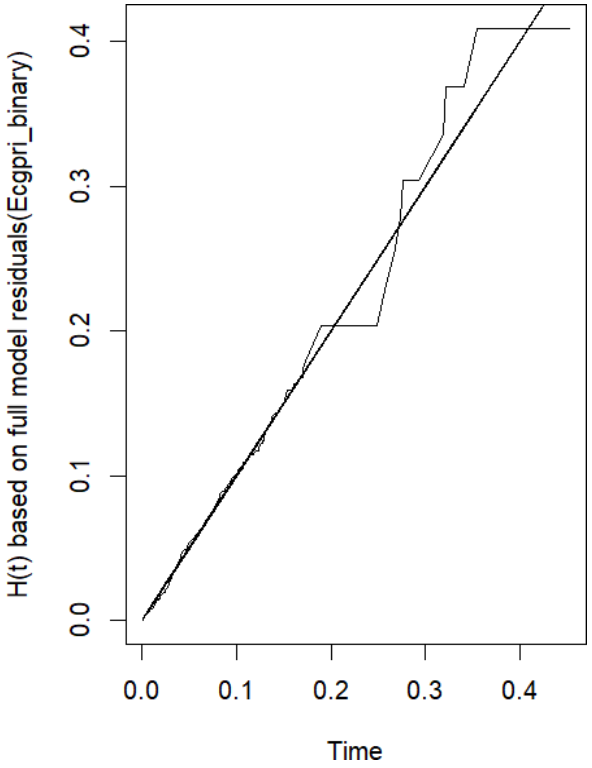

Figure 6.19: Residual plot corresponds to extended Cox model ECGpri

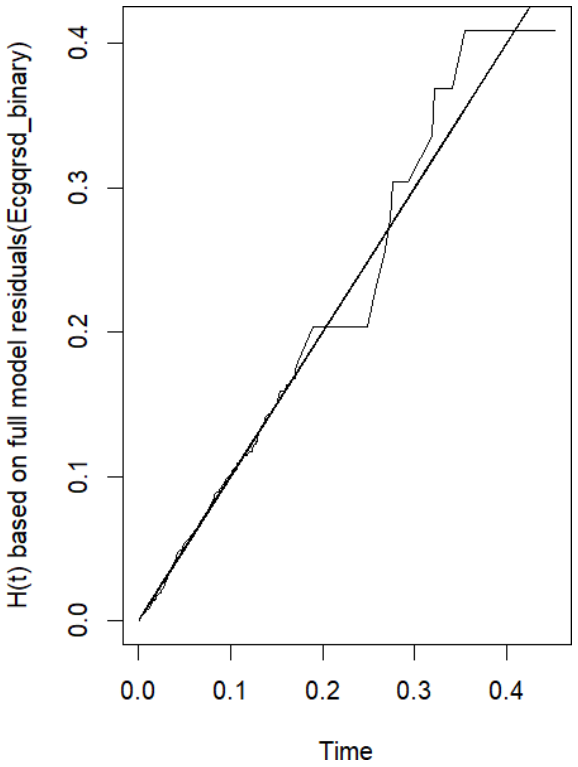

Figure 6.20: Residual plot corresponds to extended Cox model ECGqrsd

6.3.4.3 Constriction event Cox-Snell residual plots for extended Cox model

Figures 6.21 to 6.24 show the Cox-Snell residuals corresponding to the four binary ECG characteristics when looking at the constriction event outcome. The line plots slightly deviate from the reference line, which means that the extended Cox model provides a reasonably good fit.

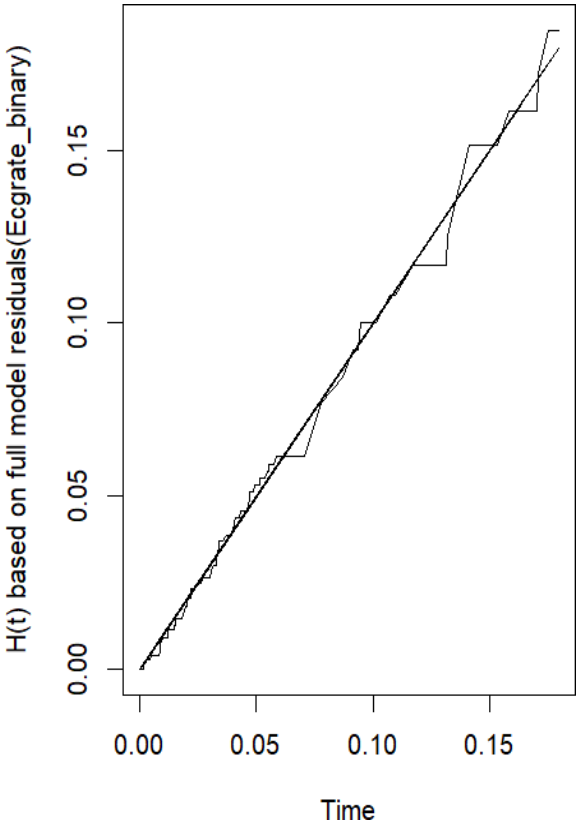

Figure 6.21: Residual plot corresponds to extended Cox model ECGrate

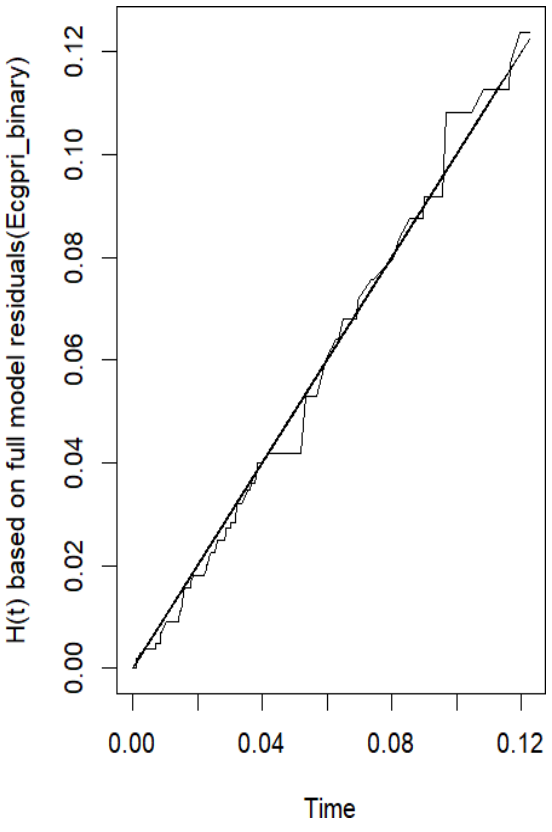

Figure 6.22: Residual plot corresponds to extended Cox model ECGqtc

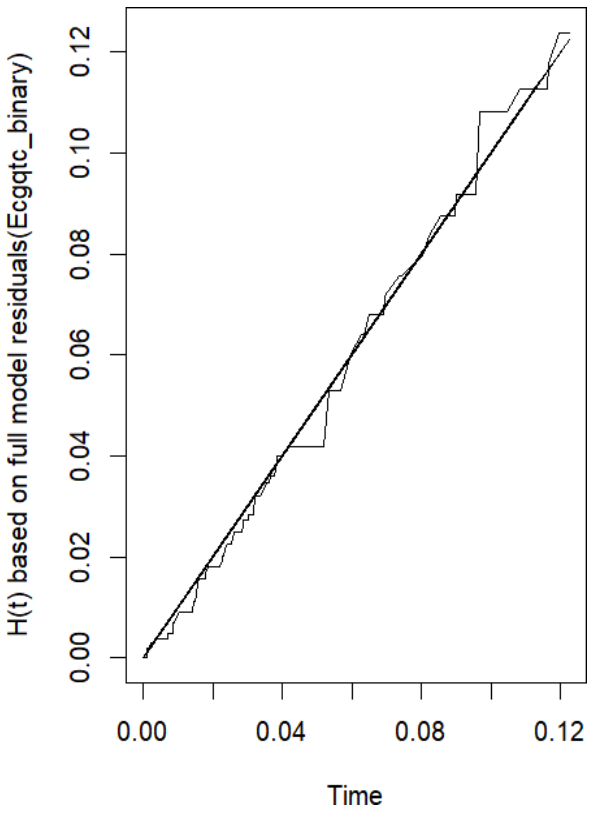

Figure 6.23: Residual plot corresponds to extended Cox model ECGpri

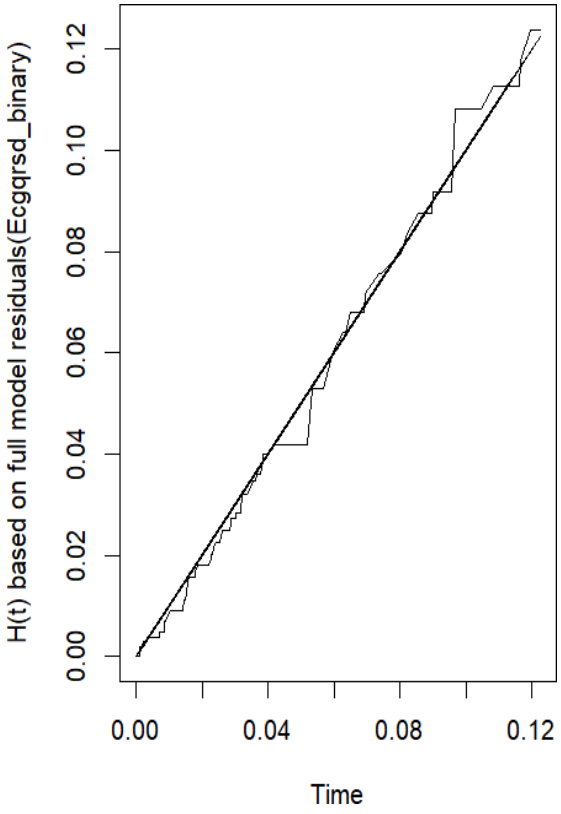

Figure 6.24: Residual plot corresponds to extended Cox model ECGqrsd

6.4 THE JOINT MODELLING OF THE IMPI TRIAL DATASET

6.4.1 The two-stage joint model

As alluded to in Chapter 4, the two-stage model was fitted in three steps: the GLMM was fitted first, then subject-specific predictions were extracted from the fitted GLMM, which were used as time-dependent covariates in the extended survival model. The GLMM was fitted using the LME4 Package in R, while the extended Cox model was implemented using the Survival Package. Table 6.7 provides the estimation results of the four binary time-varying covariates, i.e., ECGrate, ECGpri, ECGqrsd, and ECGqtc, fitted one at a time to allow for comparison with other joint models. The GLMMs were fitted using an ML based on the Laplace approximation approach, and the results are listed in Tables 6.7 to 6.10. The GLMM longitudinal sub-model of ECGrate in Table 6.7 suggests that patients on prednisolone treatment have an estimate of -0.063, which means that the mean of patients on prednisolone treatment is 0.063 less compared to patients in the placebo group for a composite event with a significant p-value of (0.001), which is evident in the standard error of 0.999, which is slightly different in the prednisolone

treatment in all three event outcomes. Furthermore, a one-unit increase in the time resulted in a 0.029 decrease in the composite outcome.

The survival sub-model in Table 6.7 shows that the duration of symptoms (days), systolic blood pressure (≥90), heart rate (≥100), definite TB pericarditis status (yes), and peripheral oedema (yes) show a significant association with the composite event, while creatanine (≥105), systolic blood pressure (≥90), and peripheral oedema (yes) show a significant association with the death event. Additionally, the duration of symptoms (days), pericardiocentesis at randomisation (yes), definite TB pericarditis status (yes), and HIV status (HIV+) also show a significant association with the composite event. There is an association between the risk of constriction and ECGrate; however, there is an increased risk of 2.219-fold in the constriction event for patients with a normal ECGrate compared to those with an abnormal ECGrate. There is also a 40% increased risk of death for patients with a normal ECGrate compared to those with an abnormal ECGrate. Moreover, there is a 54% reduced risk of the composite event for patients with a normal ECGrate compared to patients with an abnormal ECGrate. There is an 18% reduced risk of composite event and a significantly reduced risk of 59% constriction for patients on prednisolone compared to those in the placebo group, respectively. At the same time, there is a 3% increased risk of death for patients on prednisolone compared to the placebo group. It is worth noting that there is a marked difference between the ECG characteristics and prednisolone that could be the result of differences in ECG measurements taken or simply the effect of confounding variables that were fitted in the two-stage model, or even the treatment effect.

Table 6.7: Two-stage joint model parameter estimates for ECGrate

Variable	Time to composite Estimate (SE*)	p-value	Time to death Estimate (SE)	p-value	Time to constriction Estimate (SE)	p-value
Longitudinal sub-model						
Intercept	0.696 (0.999)	<0.001	0.698 (0.10)	<0.001	0.694 (0.099)	<0.001
Prednisolone	-0.063 (1.131)	0.63	-0.084 (0.131)	0.522	-0.064 (0.130)	0.622
Time	-0.029 (0.004)	<0.001	-0.026 (0.003)	<0.001	-0.030 (0.004)	<0.001

Variable	Time to composite HR (95% CI)	SE	p-value	Time to death HR (95% CI)	SE	p-value	Time to constriction HR (95% CI)	SE	p-value
Survival sub-model									
ECGrate	0.465 (0.240 - 19.314)	1.147	0.494	1.402 (0.156 - 12.581)	1.119	0.762	2.219 (1.727 - 2.535)	3.942	0.031
Prednisolone	0.821 (0.617 - 1.094)	0.147	0.178	1.025 (0.704 - 1.495)	0.192	0.894	0.410 (0.294 - 0.819)	2.292	0.008
Pericardiocentesis at randomisation (yes)	0.783 (0.535 - 1.145)	0.194	0.207	0.662 (0.409 - 1.075)	0.247	0.095	2.555 (0.178 - 3.159)	1.475	0.019
Duration of symptoms (days)	1.004 (1.001 - 1.009)	0.002	0.009	0.999 (0.994 - 1.006)	0.003	0.922	1.008 (1.001 - 1.015)	0.177	0.028
Age (years)	0.993 (0.992 - 1.015)	0.006	0.572	1.010 (0.994 - 1.028)	0.008	0.198	1.005 (0.976 - 1.027)	0.173	0.936
NYHA class at study entry:									
II	1.327 (0.835 - 2.109)	0.232	0.232	1.211 (0.653 - 2.248)	0.315	0.543	2.490 (0.518 - 4.205)	1.964	0.465
III	1.874 (1.102 - 3.186)	0.020	0.020	1.735 (0.872- 3.455)	0.351	0.116	2.001 (1.252 - 3.360)	2.149	0.019
IV	2.577 (1.370 - 4.847)	0.003	0.003	3.096 (1.442 - 6.650)	0.340	0.004	1.488 (0.520 - 3.194)	2.962	0.255
Creatinine (≥105)	1.678 (0.870 - 3.236)	0.122	0.122	2.176 (1.295 - 3.660)	0.265	0.003	2.383 (0.653 - 3.647)	7.538	0.322
HIV status (HIV+)	0.754 (0.537 - 1.061)	0.174	0.105	1.108 (0.680 - 1.808)	0.250	0.679	2.575 (0.189 - 2.941)	1.355	0.010
Weight (kg)	0.938 (0.828 - 1.064)	0.064	0.323	0.865 (0.733 - 1.021)	0.085	0.087	1.065 (0.830 - 1.367)	2.187	0.620
Palpable pulsus paradoxus (yes)	0.994 (0.681 - 1.489)	0.120	0.974	0.887 (0.541 - 1.457)	0.253	0.638	1.009 (0.459 - 2.219)	6.906	0.982
Systolic blood pressure (≥90)	0.597 (0.376 - 0.950)	0.237	0.030	0.398 (0.227 - 0.700)	0.287	0.001	1.314 (0.372 - 5.733)	1.030	0.318
Heart rate (≥100)	0.094 (0.012 - 0.746)	1.054	0.025	0.947 (0.642 - 1.399)	0.199	0.785	1.018 (0.518 - 2.000)	1.267	0.958
Effusion size:									
Medium (1-2 cm)	0.700 (0.551 - 3.706)	0.486	0.463	1.139 (0.699 - 1.856)	0.249	0.601	1.103 (0.487 - 2.496)	1.533	0.814
Small (<1 cm)	1.559 (0.211 - 1.954)	0.568	0.435	0.598 (0.245 - 1.464)	0.456	0.260	1.940 (0.113 - 2.358)	2.854	0.393
Chest X-ray pulmonary infiltrate (yes)	0.973 (0.606 - 1.743)	0.270	0.920	0.974 (0.653 - 1.454)	0.204	0.898	1.316 (0.709- 2.444)	1.162	0.385
Definite TB pericarditis status (yes)	1.517 (1.039 - 2.213)	0.193	0.031	1.229 (0.738 - 2.051)	0.261	0.428	2.781 (1.264 - 6.118)	1.480	0.011
Haemoglobin (≥10)	1.778 (0.076 - 4.242)	1.031	0.577	0.955 (0.630 - 1.449)	0.212	0.831	2.115 (0.863 - 3.503)	1.314	0.121
Peripheral oedema (yes)	1.529 (1.123 - 2.081)	0.157	0.007	1.514 (1.012 - 2.267)	0.206	0.044	1.390 (0.709 - 2.725)	1.263	0.337

* SE = standard error

Table 6.8 suggests an increased risk of a composite event by a factor of 2.053 for patients with a normal ECGpri compared to those with an abnormal ECGpri. Moreover, there is also a 28% increased risk of death for patients with a normal ECGpri compared to those with an abnormal ECGpri. In the constriction event, there is a reduced risk of 36% for normal ECGpri patients compared to abnormal ECGpri patients. There is a 19% reduced risk of the composite event for patients on prednisolone compared to the placebo group. Consequently, there is a 2% increased risk of death for patients on prednisolone compared to the placebo group. There is a 58% significantly reduced risk of constriction for patients on prednisolone compared to the placebo group. It is worth noting that there is a marked difference between the ECG characteristics and prednisolone that could be the result of differences in ECG measurements taken or simply the effect of confounding variables that were fitted in the two-stage model, or even the treatment effect.

Table 6.8: Two-stage joint model parameter estimates for ECGpri

Variable	Time to composite Estimate (SE)	p-value	Time to death Estimate (SE)	p-value	Time to constriction Estimate (SE)	p-value
Longitudinal sub-model						
Intercept	-11.572 (1.273)	<0.001	-11.641 (1.276)	<0.001	-10.730 (0.001)	<0.001
Prednisolone	-0.012 (1.456)	-0.081	-0.097 (1.462)	0.947	-11.479 (0.001)	<0.001
Time	0.003 (0.009)	0.767	0.002 (0.009)	0.779	-0.003 (0.001)	0.031

Variable	Time to composite HR (95% CI)	SE	p-value	Time to death HR (95% CI)	SE	p-value	Time to constriction HR (95% CI)	SE	p-value
Survival sub-model									
ECGpri	2.053 (0.955 - 4.415)	0.391	0.065	1.276 (0.374 - 4.357)	0.626	0.697	0.645 (0.063 - 6.583)	1.185	0.713
Prednisolone	0.812 (0.611 - 1.081)	0.146	0.154	1.019 (0.701 - 1.484)	0.192	0.918	0.420 (0.215 - 0.826)	0.344	0.012
Pericardiocentesis at randomisation (yes)	0.780 (0.533 - 1.141)	0.194	0.201	0.665 (0.410 - 1.079)	0.250	0.247	0.432 (0.196 - 0.954)	0.403	0.038
Duration of symptoms (days)	1.005 (1.001 - 1.009)	0.002	0.008	0.999 (0.994 - 1.006)	0.003	0.924	1.008 (1.001 - 1.016)	0.004	0.023
Age (years)	1.003 (0.992 - 1.015)	0.006	0.574	1.010 (0.994 - 1.028)	0.008	0.200	1.003 (0.978 - 1.029)	0.013	0.795
NYHA class at study entry:									
II	1.320 (0.830 - 2.099)	0.237	0.240	1.213 (0.654 - 2.253)	0.316	0.539	1.468 (0.520 - 4.151)	0.530	0.468
III	1.815 (1.067 - 3.086)	0.271	0.028	1.721 (0.864 - 3.429)	0.352	0.122	3.613 (1.166 - 11.204)	0.577	0.026
IV	2.521 (1.338 - 4.751)	0.323	0.004	3.079 (1.434 - 6.614)	0.390	0.004	2.428 (0.516 - 11.437)	0.791	0.262
Creatinine (≥105)	1.628 (0.847 - 3.131)	0.334	0.144	2.183 (1.298 - 3.671)	0.265	0.003	1.505 (0.639 - 3.546)	0.437	0.349
HIV status (HIV+)	0.751 (0.534 - 1.057)	0.174	0.101	1.103 (0.677 - 1.801)	0.250	0.693	0.395 (0.192 - 0.817)	0.370	0.012
Weight (kg)	0.934 (0.824 - 1.059)	0.064	0.287	0.863 (0.731 - 1.021)	0.085	0.085	1.066 (0.830 - 1.372)	0.128	0.614
Palpable pulsus paradoxus (yes)	1.006 (0.681 - 1.487)	0.199	0.976	0.881 (0.538 - 1.445)	0.252	0.617	1.056 (0.481 - 2.324)	0.402	0.891
Systolic blood pressure (≥90)	0.591 (0.372 - 0.941)	0.237	0.027	0.394 (0.225 - 0.692)	0.287	0.001	2.951 (0.390 - 22.329)	1.032	0.295
Heart rate (≥100)	0.091 (0.012 - 0.717)	1.054	0.023	0.954 (0.646 - 1.450)	0.199	0.814	1.275 (0.666 - 2.445)	0.332	0.463
Effusion size:									
Medium (1-2 cm)	1.389 (0.535 - 3.607)	0.487	0.499	1.136 (0.698 - 1.851)	0.249	0.607	1.143 (0.503 - 2.602)	0.419	0.749
Small (<1 cm)	0.617 (0.202 - 1.881)	0.569	0.396	0.591 (0.242 - 1.447)	0.456	0.250	0.504 (0.111 - 2.296)	0.773	0.376
Chest X-ray pulmonary infiltrate (yes)	1.025 (0.605 - 1.739)	0.269	0.924	0.974 (0.653 - 1.454)	0.204	0.898	1.382 (0.744- 2.568)	0.316	0.306
Definite TB pericarditis status (yes)	1.521 (1.043 - 2.220)	0.193	0.029	1.224 (0.734 - 2.043)	0.261	0.438	2.942 (1.341 - 6.454)	0.401	0.007
Haemoglobin (≥10)	0.557 (0.074 - 4.209)	1.031	0.571	0.956 (0.630 - 1.451)	0.216	0.833	1.719 (0.851 - 3.475)	0.359	0.131
Peripheral oedema (yes)	1.539 (1.131 - 2.096)	0.157	0.006	1.517 (1.014 - 2.270)	0.206	0.043	1.441 (0.736 - 2.823)	0.343	0.286

Table 6.9 suggests that there is an increased risk of a composite event by a factor of 1.403, an increased risk of death by a factor of 1.624, and an increased risk of constriction by a factor of 2.862 for all patients with a normal ECGqrsd compared to those with an abnormal ECGqrsd. At the same time, there is a 20% reduced risk of a composite event, and a significant 59% reduced risk of constriction for patients on prednisolone compared to those in the placebo group. Finally, there is a 0.8% increased risk of death for patients on prednisolone compared to the placebo group. The marked difference between the ECG characteristics and prednisolone is evident and could result from differences in ECG measurements taken or simply the effect of confounding variables that were fitted in the two-stage model, or even the treatment effect.

Table 6.9: Two-stage joint model parameter estimates for ECGqrsd

Variable	Time to composite Estimate (SE)	p-value	Time to death Estimate (SE)	p-value	Time to constriction Estimate (SE)	p-value
Longitudinal sub-model						
Intercept	-13.043 (1.891)	<0.001	-12.687 (1.789)	<0.001	-13.056 (1.896)	<0.001
Prednisolone	0.520 (2.031)	0.798	0.693 (1.877)	0.712	0.527 (2.033)	0.795
Time	0.011 (0.008)	0.172	0.011 (0.007)	0.112	0.011 (0.008)	

Variable	Time to composite HR (95% CI)	SE	p-value	Time to death HR (95% CI)	SE	p-value	Time to constriction HR (95% CI)	SE	p-value
Survival sub-model									
ECGqrsd	1.403 (0.482 - 4.085)	0.545	0.534	1.624 (0.400 - 6.599)	0.715	0.498	2.862 (0.318 - 25.756)	1.121	0.348
Prednisolone	0.805 (0.605 - 1.073)	0.146	0.140	1.008 (0.692 - 1.470)	0.192	0.965	0.407 (0.206 - 0.807)	0.348	0.010
Pericardiocentesis at randomisation (yes)	0.781 (0.534 - 1.144)	0.194	0.205	0.664 (0.409 - 1.077)	0.247	0.097	0.418 (0.190 - 0.920)	0.402	0.030
Duration of symptoms (days)	1.005 (1.001 - 1.009)	0.002	0.009	0.999 (0.994 - 1.006)	0.003	0.933	1.008 (1.001 - 1.016)	0.004	0.020
Age (years)	1.003 (0.992 - 1.016)	0.006	0.561	1.011 (0.995 - 1.028)	0.008	0.192	1.003 (0.979 - 1.029)	0.013	0.789
NYHA class at study entry:									
II	1.332 (0.839 - 2.118)	0.236	0.224	1.211 (0.653 - 2.248)	0.315	0.544	1.487 (0.525 - 4.209)	0.531	0.446
III	1.875 (1.104 - 3.188)	0.271	0.020	1.728 (0.869 - 3.439)	0.351	0.119	3.644 (1.176 - 11.299)	0.578	0.025
IV	2.559 (1.361 - 4.815)	0.322	0.004	3.046 (1.419 - 6.700)	0.400	0.004	2.369 (0.503 - 11.173)	0.791	0.275
Creatinine (≥105)	1.661 (0.863 - 3.195)	0.334	0.129	2.171 (1.290 - 3.653)	0.265	0.003	1.498 (0.637 - 3.527)	0.437	0.354
HIV status (HIV+)	0.755 (0.537 - 1.063)	0.174	0.108	1.114 (0.683 - 1.820)	0.250	0.664	0.396 (0.192 - 0.820)	0.370	0.013
Weight (kg)	0.938 (0.827 - 1.064)	0.064	0.320	0.865 (0.733 - 1.023)	0.085	0.090	1.063 (0.827 - 1.367)	0.128	0.633
Palpable pulsus paradoxus (yes)	0.996 (0.674 - 1.474)	0.200	0.987	0.883 (0.539 - 1.448)	0.252	0.622	1.063 (0.487 - 2.319)	0.398	0.878
Systolic blood pressure (≥90)	0.593 (0.373 - 0.944)	0.237	0.028	0.390 (0.223 - 0.687)	0.287	0.001	2.902 (0.384 - 21.945)	1.032	0.301
Heart rate (≥100)	0.093 (0.012 - 0.737)	1.054	0.023	0.950 (0.644 - 1.403)	0.199	0.797	1.292 (0.676 - 2.319)	0.331	0.438
Effusion size:									
Medium (1-2 cm)	1.421 (0.548 - 3.689)	0.486	0.469	1.137 (0.698 - 1.854)	0.249	0.605	1.106 (0.490 - 2.497)	0.415	0.808
Small (<1 cm)	0.635 (0.208 - 1.936)	0.568	0.425	0.597 (0.245 - 1.461)	0.456	0.259	0.491 (0.108 - 2.232)	0.772	0.357
Chest X-ray pulmonary infiltrate (yes)	1.024 (0.604 - 1.738)	0.269	0.928	0.977 (0.654 - 1.461)	0.205	0.912	1.398 (0.751 - 2.604)	0.317	0.291
Definite TB pericarditis status (yes)	1.533 (1.050 - 2.239)	0.193	0.027	1.241 (0.744 - 2.072)	0.261	0.409	3.047 (1.382 - 6.719)	0.403	0.006
Haemoglobin (≥10)	0.557 (0.074 - 4.207)	1.031	0.571	0.951 (0.628 - 1.441)	0.212	0.813	1.716 (0.850 - 3.467)	0.359	0.132
Peripheral oedema (yes)	1.529 (1.123 - 2.083)	0.157	0.007	1.519 (1.015 - 2.276)	0.206	0.042	1.410 (0.718 - 2.772)	0.345	0.319

Table 6.10 suggests that there is an increased risk of 42% of a composite event, a 16% increased risk of death, and a further increased risk of 14% in the constriction event for all patients with a normal ECGqtc compared to those with an abnormal ECGqtc. In comparison, there is a 19% reduced risk of a composite event and a significant 58% reduced risk of constriction for patients on prednisolone compared to the placebo group. In addition, there is a 2% increased risk of death for patients on prednisolone compared to the placebo group. The evident marked difference between the ECG characteristics and prednisolone could be the result of differences in ECG measurements taken or simply the effect of confounding variables that were fitted in the two-stage model, or even the treatment effect.

Table 6.10: Two-stage joint model parameter estimates for ECGqtc

Variable	Time to composite Estimate (SE)	p-value	Time to death Estimate (SE)	p-value	Time to constriction Estimate (SE)	p-value
Longitudinal sub-model						
Intercept	-10.655 (0.690)	<0.001	-10.449 (0.664)	<0.001	-10.665 (0.689)	<0.001
Prednisolone	0.101 (0.740)	0.892	0.105 (0.712)	0.883	0.101 (0.740)	0.891
Time	0.003 (0.005)	0.615	0.003 (0.005)	0.574	0.003 (0.005)	0.612

Variable	Time to composite HR (95% CI)	SE	p-value	Time to death HR (95% CI)	SE	p-value	Time to constriction HR (95% CI)	SE	p-value
Survival sub-model									
ECGqtc	1.422 (0.989 - 2.046)	0.186	0.578	1.161 (0.689 - 1.958)	0.266	0.574	1.140 (0.499 - 2.605)	0.422	0.756
Prednisolone	0.808 (0.608 - 1.076)	0.146	0.144	1.021 (0.702 - 1.486)	0.191	0.912	0.424 (0.217 - 0.830)	3.343	0.012
Pericardiocentesis at randomisation (yes)	0.804 (0.549 - 1.179)	0.195	0.264	0.669 (0.413 - 1.087)	0.247	0.105	0.426 (0.194 - 0.937)	0.401	0.034
Duration of symptoms (days)	1.004 (1.001 - 1.009)	0.002	0.009	0.999 (0.994 - 1.006)	0.003	0.934	1.008 (1.001 - 1.016)	0.004	0.022
Age (years)	1.002 (0.991 - 1.015)	0.006	0.672	1.010 (0.994 - 1.028)	0.008	0.208	1.002 (0.978 - 1.028)	0.013	0.821
NYHA class at study entry:									
II	1.346 (0.848 - 2.139)	0.236	0.207	1.223 (0.659 - 2.273)	0.316	0.523	1.483 (0.525 - 4.191)	0.530	0.457
III	1.869 (1.100 - 3.178)	0.271	0.021	1.732 (0.870 - 3.449)	0.351	0.118	3.581 (1.157 - 11.090)	0.577	0.027
IV	2.542 (1.352 - 4.782)	0.322	0.004	3.067 (1.429 - 6.588)	0.390	0.004	2.374 (0.504 - 11.209)	0.792	0.275
Creatinine (≥105)	1.608 (0.834 - 3.104)	0.335	0.156	2.169 (1.291 - 3.647)	0.265	0.003	1.490 (0.630 - 3.527)	0.439	0.364
HIV status (HIV+)	0.759 (0.540 - 1.069)	0.174	0.115	1.117 (0.684 - 1.825)	0.250	0.658	0.392 (0.193 - 0.811)	0.370	0.012
Weight (kg)	0.942 (0.831 - 1.068)	0.064	0.353	0.866 (0.733 - 1.023)	0.085	0.091	1.065 (0.829 - 1.370)	0.128	0.620
Palpable pulsus paradoxus (yes)	0.994 (0.672 - 1.471)	0.200	0.978	0.887 (0.542 - 1.457)	0.252	0.639	1.079 (0.495 - 2.353)	0.398	0.848
Systolic blood pressure (≥90)	0.593 (0.372 - 0.944)	0.237	0.027	0.394 (0.225 - 0.691)	0.287	0.001	2.970 (0.395 - 22.481)	1.032	0.292
Heart rate (≥100)	0.084 (0.011 - 0.668)	1.056	0.019	0.946 (0.641 - 1.398)	0.200	0.781	1.281 (0.669 - 2.458)	0.332	0.455
Effusion size:									
Medium (1-2 cm)	1.514 (0.580 - 3.951)	0.489	0.397	1.139 (0.699 - 1.857)	0.249	0.601	1.120 (0.497 - 2.527)	0.415	0.784
Small (<1 cm)	0.647 (0.212 - 1.978)	0.569	0.446	0.588 (0.240 - 1.441)	0.457	0.246	0.495 (0.109 - 2.255)	0.774	0.363
Chest X-ray pulmonary infiltrate (yes)	1.012 (0.595 - 1.723)	0.271	0.962	0.960 (0.642 - 1.439)	0.206	0.846	1.366 (0.735 - 2.541)	0.316	0.324
Definite TB pericarditis status (yes)	1.485 (1.017 - 2.170)	0.193	0.041	1.214 (0.726 - 2.030)	0.262	0.459	2.961 (1.349 - 6.503)	0.402	0.007
Haemoglobin (≥10)	0.582 (0.077 - 4.421)	1.034	0.601	0.947 (0.625 - 1.437)	0.212	0.801	1.718 (0.849 - 3.476)	0.359	0.132
Peripheral oedema (yes)	1.497 (1.099 - 2.042)	0.158	0.011	1.509 (1.008 - 2.261)	0.206	0.046	1.426 (0.727 - 2.801)	0.344	0.302

6.4.2 Univariate joint model

In this subsection, the four univariate shared parameter joint models were fitted for time-to-composite, time-to-death, and time-to-constriction events. The univariate joint models were fitted for four binary time-varying covariates, namely ECGrate, ECGpri, ECGqrsd, and ECGqtc, to allow for comparison with other joint models. The ECG HRs are association parameters that link ECG measurements to the risk for constriction, death, and composite, while prednisone is the effect on the active arm compared to the placebo. The univariate analysis results are presented in Tables 6.11 to 6.14. The results suggest no individual association between the ECGrate and the risk of composite, death, and constriction event outcomes. Furthermore, there is a reduced risk of 7% in the composite event and a reduced risk of 76% in the constriction event for patients with a normal ECGrate compared to those with an abnormal ECGrate. Moreover, there is an increased risk of 92% in the death event for patients with a normal ECGrate compared to the ones with an abnormal ECGrate. There is a longitudinal association between ECGpri and the risk of both death and constriction but there is no association with the composite event.

In addition, there is an increased risk of 2.077-fold in the composite event, and a reduced risk of 48% in both the death and constriction events for patients with a normal ECGpri compared to the ones with an abnormal ECGpri. There is no association between ECGqrsd and the risk of composite and death events; however, there is an association with constriction. There is a reduced risk of 64% in the constriction event, an increased risk of 2.409-fold in the composite event, and an increased risk of 7% in the death event for all patients with a normal ECGqrsd compared to those with an abnormal ECGqrsd. Finally, there is no association between ECGqtc and the risk of composite and death events; however, there is an association with constriction. There is also an increased risk of 7% in the death event, a reduced risk of 78% in the composite event, and a reduced risk of 87% in the constriction event for all patients who have a normal ECGqtc compared to those with an abnormal ECGqtc.

Table 6.11 shows a 12% reduced risk of a composite event and an 87% reduced risk of constriction for patients on prednisolone compared to the placebo group. At the same time, there is a 31% increased risk of death for patients on prednisolone compared to the placebo group.

Table 6.12 suggests that there is a 98% reduced risk of death and a 39% reduced risk of constriction for patients on prednisolone compared to the placebo group. In comparison, there is a 20% increased risk of a composite event for patients on prednisolone compared to those on the placebo.

Table 6.13 suggests that there is an 89% reduced risk of a composite event and an 83% reduced risk of constriction for patients on prednisolone compared to those on the placebo. There is, however, a 2% increased risk of death for patients on prednisolone compared to the placebo group.

Table 6.14 suggests that there is a 97% reduced risk of a composite event and a 32% reduced risk of constriction for patients on prednisolone compared to those on the placebo, respectively. Furthermore, there is a 2% increased risk of death for patients on prednisolone compared to the placebo group. Notably, there are marked differences between ECG characteristics and prednisolone, which could result from differences in ECG measurements or the effect of confounding variables that were fitted in the univariate joint model, or even the treatment effect.

Table 6.11: Univariate joint model results for ECGrate

Variable	Time to composite PM* (95% CI)	SD**	p-value	Time to death PM (95% CI)	SD	p-value	Time to constriction PM (95% CI)	SD	p-value
Longitudinal sub-model: ECGrate									
Intercept	0.731 (-0.120 - 1.621)	0.417	0.097	0.927 (0.262 - 1.649)	0.348	<0.001	-0.959 (-3.048 - 0.629)	0.980	0.307
Prednisolone	-0.826 (-2.121 - 0.0230)	0.523	0.055	-0.302 (-1.111 - 0.479)	0.406	0.447	-1.062 (-2.744 - 0.162)	0.736	0.095
Time	-0.041 (-0.165 - 0.097)	0.097	0.667	-0.122 (-0.197 - -0.037)	0.038	<0.001	-0.001 (-0.042 - 0.045)	0.025	0.965
Variable	**Time to composite HR (95% CI)**	**SD**	**p-value**	**Time to death HR (95% CI)**	**SD**	**p-value**	**Time to constriction HR (95% CI)**	**SD**	**p-value**
Survival sub-model									
Association parameter (ECGrate)	0.927 (0.658 - 1.309)	0.171	0.650	1.078 (0.645 - 1.335)	0.143	0.269	0.239 (-0.051 - 2.081)	1.119	0.382
Prednisolone	0.885 (0.276 - 2.612)	0.579	0.852	1.306 (0.189 - 9.564)	0.998	0.803	0.135 (0.001 - 56.940)	3.379	0.569
Pericardiocentesis at randomisation (yes)	0.457 (0.129 - 1.540)	0.629	0.209	0.509 (0.045 - 4.943)	1.209	0.589	1.829 (0.185 - 28.474)	1.256	<0.001
Age (years)	0.987 (0.949 - 1.024)	0.019	0.488	0.915 (0.800 - 1.026)	0.064	0.148	0.786 (0.617 - 1.022)	0.126	0.063
NYHA class at study entry:									
II	0.533 (0.143 - 2.171)	0.687	0.352	2.284 (0.25 - 34.090)	1.590	0.055	2.883 (0.645 - 12.616)	0.769	<0.001
III	0.618 (0.155 - 2.678)	0.718	0.484	1.524 (0.704 - 3.357)	0.391	0.097	0.661 (0.095 - 6.318)	4.266	0.067
IV	0.361 (0.048 - 2.138)	0.953	0.277	1.179 (0.042 - 44.467)	1.703	0.921	1.225 (0.051 - 27.058)	1.583	<0.001
HIV status (HIV+)	0.512 (0.156 - 1.629)	0.588	0.246	1.183 (0.540 - 2.866)	0.406	<0.001	0.028 (0.001 - 1.291)	2.003	<0.001
Palpable pulsus paradoxus (yes)	1.533 (0.514 - 4.289)	0.536	0.403	1.819 (0.314 - 11.693)	0.912	0.497	2.351 (0.017 - 356.737)	2.474	0.719
Definite TB pericarditis status (yes)	3.804 (1.119 - 14.069)	0.645	0.036	1.252 (0.059 - 17.637)	1.435	0.828	2.293 (0.961 - 5.104)	0.430	<0.001
Peripheral oedema (yes)	3.152 (1.251 - 8.191)	0.478	0.015	0.296 (0.032 - 2.482)	1.096	0.260	1.616 (0.748 - 3.290)	0.390	<0.001

* PM = posterior mean; **SD = standard deviation

Table 6.12: Univariate joint model results for ECGpri

Variable	Time to composite PM (95% CI)	SD	p-value	Time to death PM (95% CI)	SD	p-value	Time to constriction PM (95% CI)	SD	p-value
Longitudinal sub-model: ECGpri									
Intercept	-4.078 (-6.567 - -2.670)	0.980	<0.001	-3.647 (-5.775 - -2.405)	0.826	<0.001	-4.179 (-6.678 - -3.251)	0.865	<0.001
Prednisolone	-0.413 (-1.742 - 0.952)	0.730	0.580	-0.541 (-2.217 - 0.827)	0.783	0.480	-0.218 (-0.996 - 0.596)	0.409	0.578
Time	0.007 (-0.006 - 0.019)	0.006	0.250	0.002 (-0.009 - 0.012)	0.005	0.650	0.006 (-0.002 - 0.015)	0.004	0.136
Variable	**Time to composite HR (95% CI)**	**SD**	**p-value**	**Time to death HR (95% CI)**	**SD**	**p-value**	**Time to constriction HR (95% CI)**	**SD**	**p-value**
Survival sub-model									
Association parameter (ECGpri)	2.077 (0.642 - 3.300)	0.379	0.124	0.522 (0.014 - 0.556)	0.143	<0.001	0.519 (0.432 - 3.871)	0.464	<0.001
Prednisolone	1.196 (0.008 - 26.602)	1.692	0.388	0.020 (0.001 - 1.437)	0.255	0.192	0.612 (0.284 - 3.909)	0.123	0.192
Pericardiocentesis at randomisation (yes)	0.672 (0.674 - 0.697)	0.117	0.156	0.555 (0.222 - 1.289)	0.426	0.151	0.330 (0.112 - 1.044)	0.585	0.070
Age (years)	0.924 (0.779 - 1.092)	0.070	0.184	1.007 (0.978 - 1.038)	0.014	0.536	0.960 (0.915 - 1.008)	0.026	0.136
NYHA class at study entry:									
II	0.727 (0.036 - 2.104)	2.723	0.494	1.347 (0.526 - 4.088)	0.491	0.499	2.884 (0.645 - 12.616)	0.769	0.194
III	0.009 (0.001 - 1.869)	4.118	0.276	2.527 (0.889 - 9.826)	0.596	0.081	1.833 (0.329 - 9.545)	0.857	0.456
IV	3.550 (0.009 - 816.478)	2.375	0.264	6.679 (1.631 - 63.625)	0.978	0.002	0.288 (0.001 - 10.371)	2.345	0.652
HIV status (HIV+)	0.647 (0.043 - 16.281)	1.272	0.326	1.183 (0.540 - 2.866)	0.406	0.636	0.464 (0.154 - 1.218)	0.517	0.102
Palpable pulsus paradoxus (yes)	1.659 (0.179 - 40.609)	1.175	0.262	0.701 (0.247 - 1.499)	0.463	0.424	0.985 (0.030 - 8.689)	0.800	<0.001
Definite TB pericarditis status (yes)	1.484 (0.132 - 124.089)	1.517	0.318	1.002 (0.324 - 2.438)	0.489	0.893	0.918 (0.037 - 9.098)	1.461	0.908
Peripheral oedema (yes)	0.020 (0.001 - 48.521)	3.085	0.278	1.593 (0.812 - 3.626)	0.369	0.152	0.514 (0.112 - 2.307)	0.785	0.428

Table 6.13: Univariate joint model results for ECGqrsd

Variable	Time to composite PM (95% CI)	SD	p-value	Time to death PM (95% CI)	SD	p-value	Time to constriction PM (95% CI)	SD	p-value
Longitudinal sub-model: ECGqrsd									
Intercept	-9.771 (-17.272 - -5.343)	3.014	<0.001	-7.654 (-12.111 - -5.011)	1.841	<0.001	-9.542 (-14.589 - -6.053)	2.300	<0.001
Prednisolone	0.761 (-2.234 - 3.981)	1.512	0.594	1.312 (-1.039 - 3.958)	1.277	0.294	1.053 (-0.501 - 2.830)	0.873	0.186
Time	0.021 (0.004 - 0.043)	0.010	0.022	0.015 (0.003 - 0.028)	0.006	0.010	0.016 (0.004 - 0.029)	0.007	<0.001
Variable	**Time to composite HR (95% CI)**	**SD**	**p-value**	**Time to death HR (95% CI)**	**SD**	**p-value**	**Time to constriction HR (95% CI)**	**SD**	**p-value**
Survival sub-model									
Association parameter (ECGqrsd)	2.409 (0.647 - 4.310)	0.482	0.118	1.071 (0.933 - 1.230)	0.071	0.321	0.360 (0.211 - 0.608)	0.349	0.030
Prednisolone	0.107 (0.006 - 1.615)	1.507	0.118	1.017 (0.215 - 6.160)	0.763	0.907	0.175 (0.039 - 0.855)	0.774	0.044
Pericardiocentesis at randomisation (yes)	2.268 (0.482 - 9.622)	0.779	0.304	0.555 (0.222 - 1.289)	0.426	0.151	0.315 (0.020 - 1.755)	2.678	0.686
Age (years)	0.964 (0.871 - 1.056)	0.049	0.470	1.007 (0.978 - 1.08)	0.014	0.536	1.088 (1.063 - 1.111)	0.011	<0.001
NYHA class at study entry:									
II	0.812 (0.205 - 2.884)	0.669	0.720	1.347 (0.369 - 2.737)	0.491	0.499	2.884 (0.645 - 12.616)	0.769	<0.001
III	0.119 (-1.159 - 1.352)	0.119	0.864	2.527 (0.014 - 5.897)	0.596	0.081	1.833 (0.329 - 9.545)	0.857	<0.001
IV	0.695 (0.116 - 3.572)	0.869	0.690	6.679 (1.631 - 63.625)	0.978	<0.001	5.658 (1.813 - 18.449)	0.598	<0.001
HIV status (HIV+)	0.761 (0.268 - 2.088)	0.519	0.602	1.183 (0.540 - 2.866)	0.406	0.636	0.140 (0.076 - 0.277)	0.784	<0.001
Palpable pulsus paradoxus (yes)	1.278 (1.066 - 1.511)	0.093	<0.001	0.700 (0.247 - 1.499)	0.463	0.424	0.003 (0.001 - 1.130)	3.347	0.058
Definite TB pericarditis status (yes)	0.889 (0.057 - 13.791)	1.408	0.922	1.002 (0.324 - 2.438)	0.489	0.893	1.756 (1.528 - 2.012)	0.068	<0.001
Peripheral oedema (yes)	0.868 (0.201 - 4.072)	0.745	0.778	1.594 (0.812 - 3.626)	0.369	0.152	0.972 (0.735 - 1.388)	0.154	0.772

Table 6.14: Univariate joint model results for ECGqtc

Variable	Time to composite PM (95% CI)	SD	p-value	Time to death PM (95% CI)	SD	p-value	Time to constriction PM (95% CI)	SD	p-value
Longitudinal sub-model: ECGqtc									
Intercept	-4.549 (-8.839 - -2.380)	1.765	<0.001	-2.911 (-4.208 - -1.936)	0.586	<0.001	-2.722 (-3.973 - -1.969)	0.534	<0.001
Prednisolone	0.697 (-0.710 - 2.611)	0.823	0.320	-0.022 (-1.266 - 1.079)	0.607	0.974	0.173 (-0.336 - 0.696)	0.251	0.472
Time	0.006 (-0.006 - 0.021)	0.007	0.394	0.003 (-0.004 - 0.011)	0.004	0.400	0.001 (-0.007 - 0.008)	0.004	0.906
Variable	**Time to composite HR (95% CI)**	**SD**	**p-value**	**Time to death HR (95% CI)**	**SD**	**p-value**	**Time to constriction HR (95% CI)**	**SD**	**p-value**
Survival sub-model									
Association parameter (ECGqtc)	0.203 (0.065 - 3.695)	0.804	0.066	1.071 (0.933 - 1.230)	0.071	0.321	0.129 (0.053 - 2.344)	0.846	<0.001
Prednisolone	0.028 (0.001 - 6.110)	2.675	0.196	1.017 (0.215 - 6.160)	0.763	0.907	0.685 (0.306 - 1.232)	0.308	0.126
Pericardiocentesis at randomisation (yes)	0.039 (0.020 - 0.134)	5.037	<0.001	0.555 (0.222 - 1.289)	0.426	0.151	1.647 (0.040 - 10.085)	1.191	0.314
Age (years)	0.869 (0.775 - 0.972)	0.059	0.028	1.007 (0.978 - 1.038)	0.014	0.536	0.903 (0.887 - 0.926)	0.009	<0.001
NYHA class at study entry:									
II	0,113 (0.058 - 0.281)	0.403	<0.001	1.347 (0.026 - 4.088)	0.491	0.499	2.316 (0.919 - 6.265)	0.507	0.088
III	0.316 (0.177 - 0.643)	0.316	0.004	2.527 (0.889 - 9.82)	0.596	0.081	3.931 (1.527- 10.979)	0.515	0.002
IV	0.163 (0.066 - 0.492)	0.509	0.002	6.679 (1.631 - 9.786)	0.978	0.002	2.924 (0.827 - 9.924)	0.631	0.098
HIV status (HIV+)	0.189 (0.103 - 0.432)	0.361	<0.001	1.183 (0.540 - 2.866)	0.406	0.636	0.009 (0.001 - 0.045)	1.443	0.002
Palpable pulsus paradoxus (yes)	2.158 (1.158 - 4.051)	0.309	0.020	0.700 (0.247 - 1.499)	0.463	0.424	0.836 (0.394 - 1.791)	0.364	0.290
Definite TB pericarditis status (yes)	1.323 (1.160 - 1.429)	0.054	<0.001	1.002 (0.324 - 2.438)	0.489	0.893	2.179 (1.508 - 2.519)	0.128	<0.001
Peripheral oedema (yes)	3.593 (1.647 - 6.613)	0.354	0.004	1.594 (0.812 - 3.626)	0.369	0.152	0.835 (-0.394 - 1.791)	0.364	0.290

6.4.3 Joint modelling overall model fit

This subsection shows the Cox-Snell residual plots used to assess the overall fit of the joint models. The subsection first starts with the two-stage joint modelling, which is followed by the univariate joint models, and finally the multivariate joint models for all four binary ECG characteristics looking at time-to-composite, time-to-death, and time-to-constriction event outcomes to provide insight into how well the models fit the data.

6.4.3.1 Time-to-composite event

Figures 6.25 to 6.28 show Cox-Snell residual plots of ECGrate, ECGqrsd, ECGpri, and ECGqtc looking at the composite event outcome. All four line plots deviate from the reference line, which means that the two-stage joint model is not a good fit for the IMPI data in the composite event outcome.

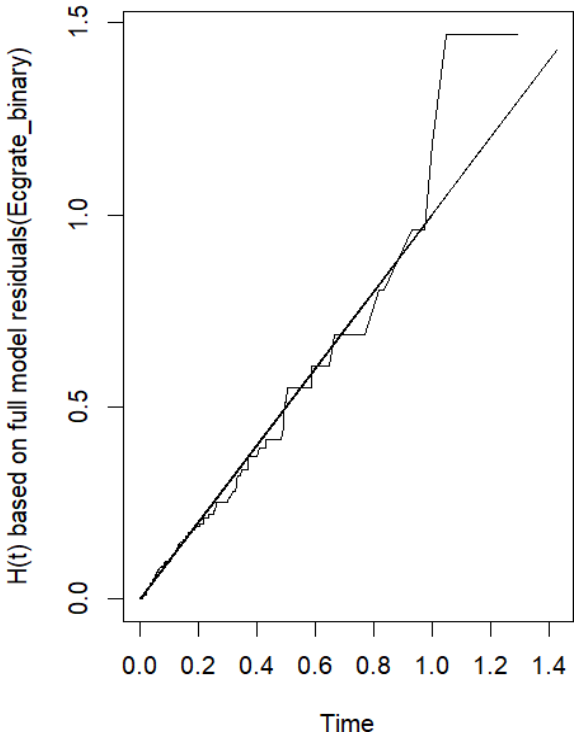

Figure 6.25: Residual plot corresponds to extended Cox model ECGrate

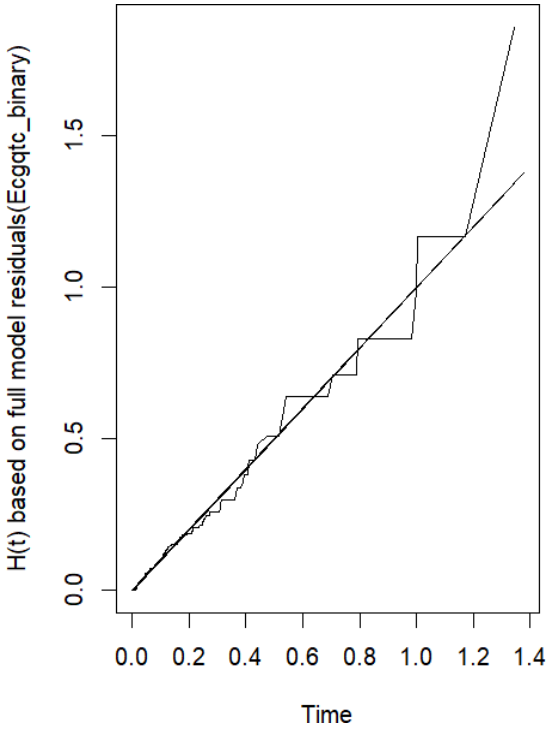

Figure 6.26: Residual plot corresponds to extended Cox model ECGqtc

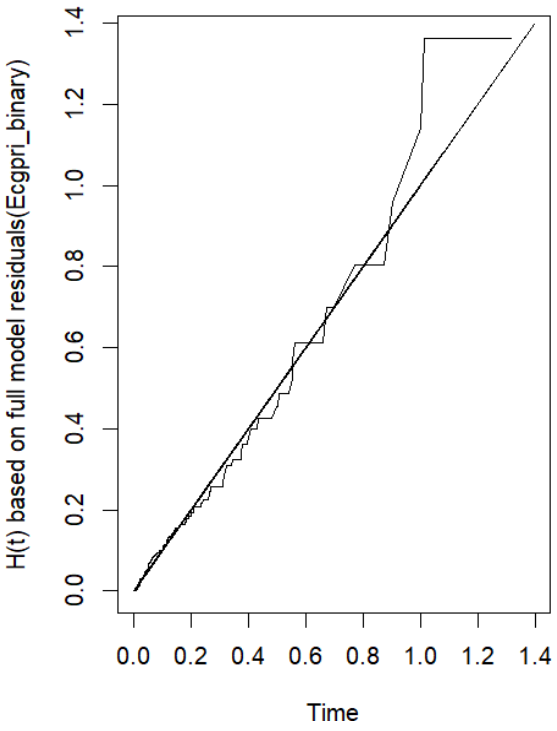

Figure 6.27: Residual plot corresponds to extended Cox model ECGpri

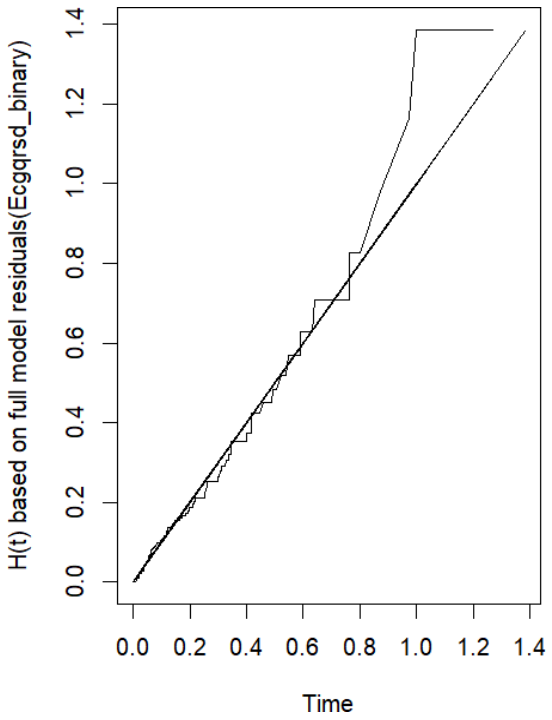

Figure 6.28: Residual plot corresponds to cxtended Cox model ECGqrsd

6.4.3.2 Time-to-death event

Figures 6.29 to 6.32 show the Cox-Snell residual plots of ECGrate, ECGqrsd, ECGpri, and ECGqtc looking at the death event outcome. All four line plots still deviate from the reference line, which means that the two-stage joint model is not a good fit for the IMPI data in the death event outcome.

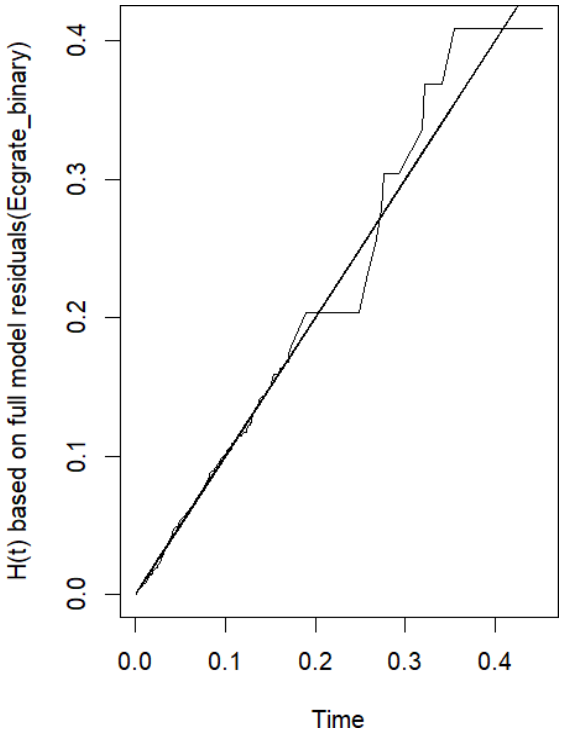

Figure 6.29: Residual plot corresponds to extended Cox model ECGrate

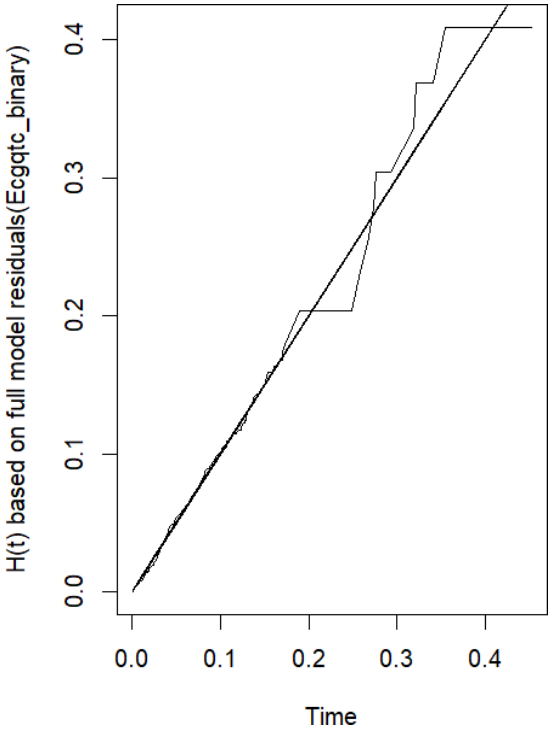

Figure 6.30: Residual plot corresponds to extended Cox model ECGqtc

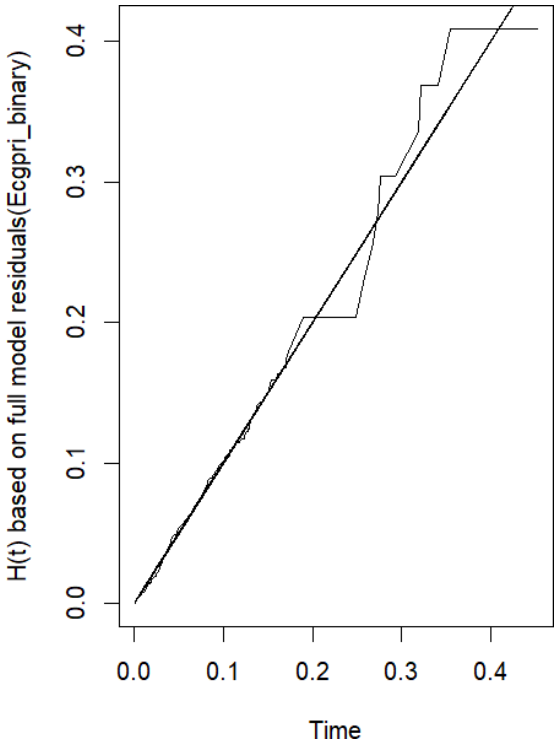

Figure 6.31: Residual plot corresponds to extended Cox model ECGpri

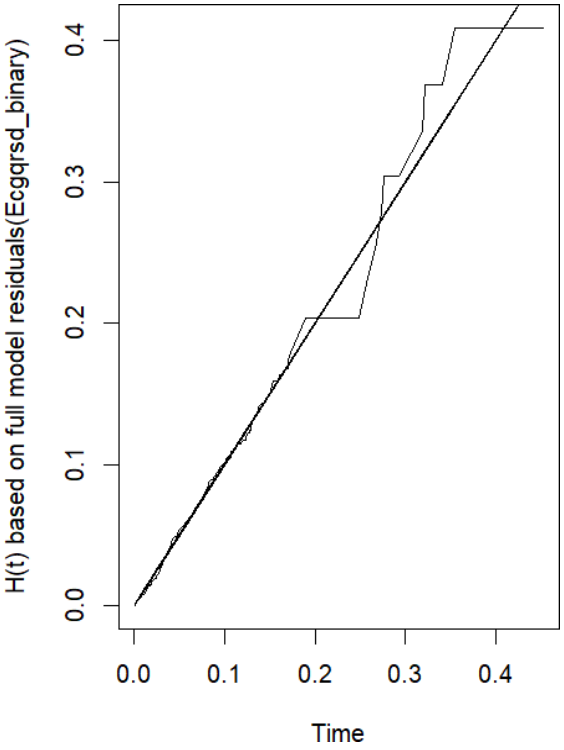

Figure 6.32: Residual plot corresponds to extended Cox model ECGqrsd

6.4.3.3 Time-to-constriction event

Figures 6.33 to 6.36 show the Cox-Snell residual plots of ECGrate, ECGqrsd, ECGpri, and ECGqtc looking at the constriction event outcome. None of the four line plots deviate much from the reference line, except for the ECGrate line plot, which suggests that the two-stage joint model is a reasonably good fit for the IMPI data in the constriction event outcome.

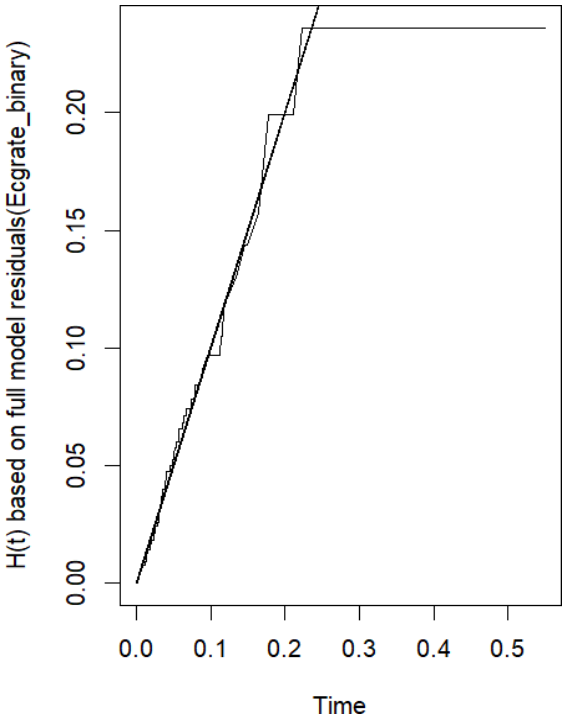

Figure 6.33: Residual plot corresponds to extended Cox model ECGrate

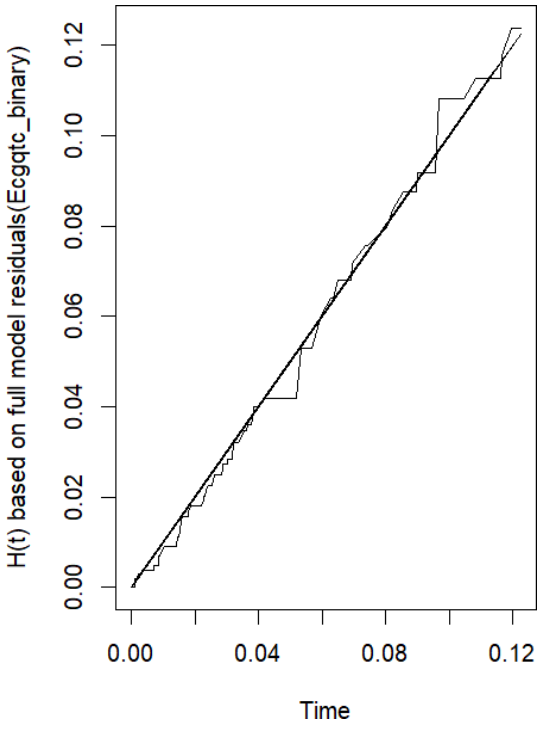

Figure 6.34: Residual plot corresponds to extended Cox model ECGqtc

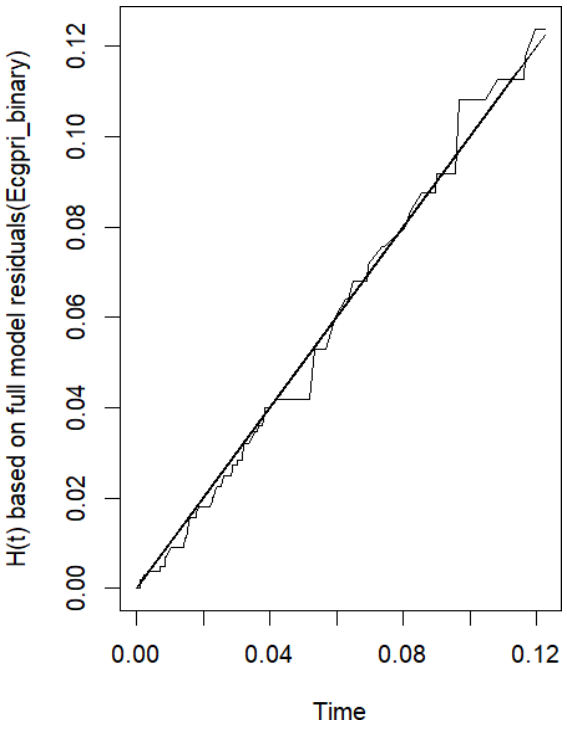

Figure 6.35: Residual plot corresponds to extended Cox model ECGpri

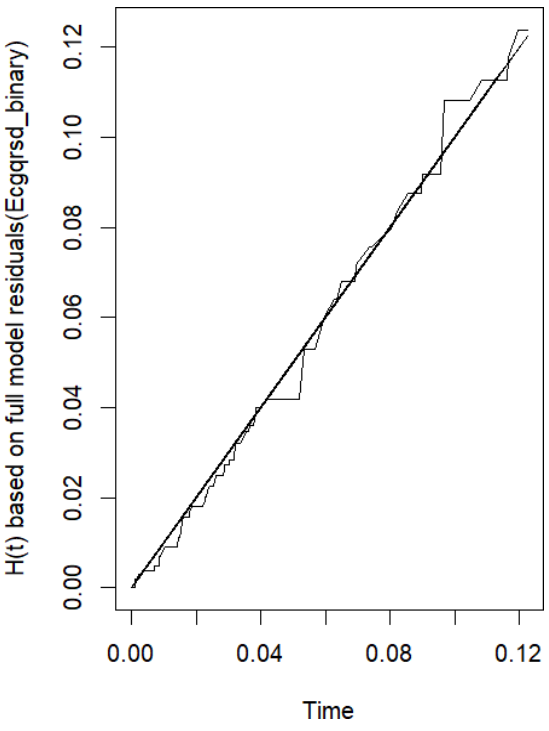

Figure 6.36: Residual plot corresponds to extended Cox model ECGqrsd

6.4.4 Shared parameter joint model

So far, model checking was done for the standard Cox PH model, the extended Cox model, and, finally, the two-stage model. This was done using the Cox-Snell residual plots, which look at four binary ECG variables for time-to-composite, time-to-death, and time-to-constriction events. A model-checking technique exists in the frequentist framework when estimation is done using ML theory through numeric integration in both the univariate and multivariate shared parameter joint model settings. However, when the estimation of parameters is done using the Bayesian approach, which is the case in the shared parameter joint models in the GLMM framework, no model-checking tools are currently developed; this is a potential area of further research.

6.4.5 Multivariate joint model for longitudinal and time-to-event data

The same GLMM as in the one used in Section 4.3 is first implemented. In this model, all subjects in the study have the same ECG characteristic trajectory in time but only differ at baseline. Each subject therefore has an intercept, which is why the random intercept-only model is used. To support this notion, it is clearly visible from Tables 6.17 to 6.19 in terms of the

Akaike information criterion (AIC) and the Bayesian information criterion (BIC) that the random intercepts-only model is the best model for ECGpri, ECGqrsd, and ECGqrsd. However, regarding ECGrate and random intercept, slope is the best model in terms of the AIC and BIC, as shown in Table 6.16. The same multivariate joint model as the one used in Section 5.2 was fitted. The ECG HRs presented in Table 6.15 are association parameters that link ECG measurements to the risk for constriction, death, or composite events. Prednisone, on the other hand, is the effect on the active arm compared to the placebo.

The results are as follows. Firstly, there is a notable slight increase in the PM of all three events from the GLMM estimates. Furthermore, there is a subject-specific effect of time for both ECGrate and ECGqrsd. Table 6.15 further shows the survival parameter estimates for time to composite, time to death, and time to constriction. There is a 9% increased risk of composite event for patients with a normal ECGrate compared to those with an abnormal ECGrate. Moreover, there is a 7% increased risk of death for patients with a normal ECGrate compared to those with an abnormal ECGrate. However, there is a 44% reduction in the risk of constriction for patients with a normal ECGrate compared to those with an abnormal ECGrate.

Patients who have a normal ECGpri rate have a higher risk of 2.130-fold for experiencing death compared to those with an abnormal ECGpri rate. There is also an 89% reduction in the risk of the composite event and a 99% reduction in the risk of constriction for patients with a normal ECGpri compared to those with an abnormal ECGpri. Additionally, patients with a normal ECGqrsd have a 5% lower risk of experiencing the composite event, a 13% lower risk of experiencing death, and a 2% lower risk of experiencing constriction. There is a higher risk of 2.026-fold for experiencing either death, constriction, or tamponade and a higher risk of 1.430-fold for experiencing constriction for patients with a normal ECGqtc compared with those with an abnormal ECGqtc. A reduction of 30% risk of death for normal ECGqtc patients compared to abnormal ECGqtc patients was also observed.

There is a 40% reduced risk of the composite event and an 88% reduced risk of constriction for patients on prednisolone compared to those in the placebo group, respectively. Moreover, there is a 2% increased risk of death for patients on prednisolone compared to those in the placebo group. The marked difference between the ECG HR and prednisolone could be as a result of differences in ECG measurements taken or the effect of confounding variables that were fitted in the multivariate joint model, or even the treatment effect.

It is worth noting that there is no joint association between the ECGrate and the risk of composite, death, and constriction event outcomes, as shown by overlapping 95% confidence intervals. Also, there is no joint association between the ECGpri and the risk of composite, death, and constriction event outcomes. There is also no joint association between the ECGqrs duration and the risk of composite, death, and constriction event outcomes. Finally, there is no joint association between ECGqtc and the risk of composite and constriction event outcomes. However, there is a joint association between ECGqtc and the risk of death.

Table 6.15: Multivariate joint model results

Variable	Time to composite PM (95% CI)	SD	p-value	Time to death PM (95% CI)	SD	p-value	Time to constriction PM (95% CI)	SD	p-value
Longitudinal sub-model: ECGrate									
Intercept	1.891 (0.601 - 5.760)	0.511	0.201	2.387 (1.680 - 3.212)	0.203	<0.001	3.999 (1.228 - 55.813)	1.020	0.027
Prednisolone	0.920 (0.715 - 1.167)	0.113	0.335	0.931 (0.709 - 1.218)	0.127	0.407	0.214 (0.004 - 0.804)	1.431	0.022
Time	0.974 (0.966 - 0.980)	0.004	<0.001	0.974 (0.967 - 0.981)	0.004	<0.001	0.881 (0.660 - 1.10)	0.170	0.667
Longitudinal sub-model: ECGpri									
Intercept	0.017 (0.010 - 0.027)	0.222	<0.001	0.015 (0.006 - 0.033)	0.399	<0.001	0.026 (0.015 - 0.039)	0.235	<0.001
Prednisolone	0.889 (0.500 - 1.735)	0.288	0.537	0.838 (0.407 - 1.645)	0.315	0.400	0.888 (0.56 - 1.639)	0.298	0.655
Time	1.003 (0.996 - 1.011)	0.004	0.522	1.005 (0.970 - 1.010)	0.003	0.206	1.008 (0.999 – 1.014)	0.004	0.055
Longitudinal sub-model: ECGqrsd									
Intercept	0.001 (0.000 - 0.002)	2.121	<0.001	0.001 (0.000 - 0.002)	1.789	<0.001	0.001 (0.000 - 0.002)	1.588	<0.001
Prednisolone	2.927 (0.949 - 7.171)	0.493	0.060	3.216 (0.918 - 11.834)	0.589	0.064	2.812 (0.786 - 12.782)	0.696	0.112
Time	1.116 (1.008 - 1.349)	0.134	<0.001	1.116 (1.008 - 1.349)	0.133	<0.001	1.017 (1.005 - 1.029)	0.006	0.008
Longitudinal sub-model: ECGqtc									
Intercept	0.119 (0.075 - 0.198)	0.273	<0.001	0.128 (0.086 - 5.244)	0.251	<0.001	0.083 (0.040 - 0.142)	0.336	<0.001
Prednisolone	1.161 (0.823 - 1.650)	0.166	0.304	1.139 (0.830 - 1.551)	0.146	0.320	1.189 (0.764 - 1.880)	0.232	0.449
Time	0.999 (0.995 - 1.004)	0.002	0.559	0.998 (0.994 - 1.003)	0.002	0.346	0.999 (0.993 - 1.006)	0.003	0.859
Variable	**Time to composite HR (95% CI)**	**SD**	**p-value**	**Time to death HR (95% CI)**	**SD**	**p-value**	**Time to constriction HR (95% CI)**	**SD**	**p-value**
Survival sub-model									
Association parameter (ECGrate)	1.087 (0.514 - 1.355)	0.091	0.321	1.071 (0.933 - 1.230)	0.071	0.321	0.559 (0.016 - 4.289)	1.779	0.689
Association parameter (ECGpri)	0.114 (0.007 - 1.138)	1.409	0.224	2.130 (0.822 - 9.924)	0.670	0.229	0.002 (0.001- 8.647)	3.786	0.104
Association parameter (ECGqrsd)	0.947 (0.715 - 1.342)	0.149	0.441	0.872 (0.621 - 1.21)	0.178	0.262	0.979 (0.352 - 2.457)	0.491	0.981
Association parameter (ECGqtc)	2.026 (0.739 - 12.061)	0.804	0.061	0.700 (0.025 - 10.881)	1.462	0.019	1.430 (0.270 - 8.174)	0.797	0.632
Prednisolone	0.599 (0.114 - 2.649)	0.746	0.403	1.017 (0.215 - 6.160)	0.763	0.907	0.124 (0.002 - 7.791)	2.082	0.280
Pericardiocentesis at randomisation (yes)	0.617 (0.296 - 1.156)	0.329	0.112	0.555 (0.222 - 1.289)	0.426	0.151	0.331 (0.107 - 0.850)	0.519	0.021
Age (years)	1.005 (0.983 - 1.025)	0.010	0.518	1.007 (0.978 - 1.038)	0.014	0.536	0.995 (0.967 - 1.023)	0.012	0.434
NYHA class at study entry:									
II	1.737 (0.868 - 3.896)	0.371	0.111	1.347 (0.526 - 4.088)	0.491	0.499	1.072 (-0.204 - 2.655)	0.724	0.096
III	3.357 (1.474 - 10.528)	0.542	0.001	2.527 (0.889 - 9.826)	0.596	0.081	5.978 (1.419 - 35.481)	0.840	0.013
IV	4.319 (1.406 - 19.609)	0.730	0.006	6.679 (1.631 - 63.625)	0.978	0.002	6.316 (1.011 - 51.316)	1.010	0.048
HIV status (HIV+)	0.521 (0.237 - 0.900)	0.358	0.015	1.183 (0.540 - 2.866)	0.406	0.636	0.192 (0.058 - 0.464)	0.528	0.001
Palpable pulsus paradoxus (yes)	1.039 (0.499 - 2.022)	0.339	0.836	0.701 (0.247 - 1.499)	0.463	0.424	0.713 (0.274 - 1.441)	0.429	0.423
Definite TB pericarditis status (yes)	2.147 (1.129 - 4.904)	0.360	0.022	1.002 (0.324 - 2.438)	0.489	0.893	5.601 (1.910 - 24.240)	0.648	0.001
Peripheral oedema (yes)	1.870 (1.121 - 4.007)	0.333	0.015	1.594 (0.812 - 3.626)	0.369	0.152	1.865 (0.829 - 4.702)	0.441	0.136

Table 6.16: GLMM results for ECGrate

Parameter	Random intercept only		Random intercept and slope	
	Estimate (SE)	p-value	Estimate (SE)	p-value
Intercept	0.826 (0.124)	0.018	0.893 (0.158)	0.018
Prednisolone	-0.156 (0.157)	0.323	-0.169 (0.153)	0.271
Time	-0.033 (0.006)	0.018	-0.095 (0.020)	0.018
Prednisolone × Time	0.009 (0.007)	0.186	0.025 (0.019)	0.170
σ_b^2	0.007	-	0.008	-
σ_c^2	0.037	-	0.033	-
Log likelihood	-736.409	-	330.249	-
AIC	1482.82	-	-314.2494	-
BIC	1507.349	-	-273.8762	-

Table 6.17: GLMM results for ECGpri

Parameter	Random intercept only		Random intercept and slope	
	Estimate (SE)	p-value	Estimate (SE)	p-value
Intercept	-3.949 (1.311)	0.003	-3.442 (0.313)	0.018
Prednisolone	-0.152 (0.392)	0.698	-0.142 (0.351)	0.686
Time	-0.004 (0.006)	0.498	0.000 (0.006)	0.652
Prednisolone × Time	0.001 (0.008)	0.879	0.005 (0.008)	0.544
σ_b^2	0.007	-	0.008	-
σ_c^2	0.037	-	0.033	-
Log likelihood	-159.478	-	-160.691	-
AIC	328.956	-	335.382	-
BIC	353.485	-	369.723	-

Table 6.18: GLMM results for ECGqrsd

Parameter	Random intercept only		Random intercept and slope	
	Estimate (SE)	p-value	Estimate (SE)	p-value
Intercept	-8.464 (1.138)	0.018	-10.689 (2.426)	0.018
Prednisolone	1.184 (0.743)	0.111	1.392 (0.932)	0.135
Time	0.021 (0.008)	0.009	0.036 (0.018)	0.039
Prednisolone × Time	-0.009 (0.011)	0.475	-0.008 (0.012)	0.475
σ_b^2	0.007	-	0.008	-
σ_c^2	0.037	-	0.033	-
Log likelihood	-120.904	-	-120.082	-
AIC	251.809	-	254.163	-
BIC	276.337	-	288.503	-

Table 6.19: GLMM results for ECGqtc

Parameter	Random intercept only		Random intercept and slope	
	Estimate (SE)	p-value	Estimate (SE)	p-value
Intercept	-2.537 (0.405)	0.018	-2.822 (0.576)	0.018
Prednisolone	0.064 (0.246)	0.795	0.054 (0.273)	0.844
Time	-0.0025 (0.005)	0.589	0.000 (0.007)	0.909
Prednisolone × Time	-0.008 (0.006)	0.228	0.007 (0.006)	0.244
σ_b^2	0.007	-	0.008	-
σ_c^2	0.037	-	0.033	-
Log likelihood	-516.395	-	-516.049	-
AIC	1042.789	-	1046.099	-
BIC	1067.318	-	1080.439	-

CHAPTER 7:
DISCUSSION, CONCLUSIONS, AND FUTURE RESEARCH

7.1 INTRODUCTION

This chapter summarises and concludes this study and suggests further research. The bibliography containing all the sources used for this study follows, and thereafter the appendices, which contain the continuous ECG results for reference (see Appendix A), as well as the R code (see Appendix B) used for the analyses in this study.

7.2 SUMMARY AND CONCLUSION

The main aim of this study was to assess the effect of the ECG characteristics, namely ECGrate, ECGpri, ECGqtc, and ECGqrsd, which are the time-varying covariates, on survival outcomes, death, pericardial constriction, and composite event (death, constriction, or cardiac tamponade, whichever happens first). This was achieved by analysing the IMPI clinical trial dataset. The analysis was restricted to patients with only ECG data and the analysis focused on patients with at least two ECG measurements. The study included 1 197 patients, of whom 845 were HIV positive and 352 were HIV negative. Five hundred and ninety-three patients were randomised to a prednisolone arm, and 600 were randomised to a placebo arm. Two hundred and fifty-five patients among the 1 197 patients experienced a composite event (i.e., death, cardiac tamponade, or constriction), 163 died, and 86 suffered cardiac arrest. Specifically, there were 71% HIV-positive and 29% HIV-negative patients. All the patients experienced a composite event, 14% died, and 7% suffered cardiac arrest.

The treatment group analysis was restricted to only prednisolone and placebo. The reason is that the comparison between MIP and the placebo was stopped early for futility in February 2013 (Mayosi et al., 2014). There were no statistically significant differences between the treatments in baseline characteristics. Notably, the ECG characteristics from the IMPI trial dataset were transformed from continuous to categorical variables to make the variables clinically meaningful. Therefore, they were explained in terms of each ECG variable's normal versus abnormal rate. The ECG characteristics were examined individually and jointly, comparing death versus constriction and composite (death, constriction, or cardiac tamponade) events. The standard Cox model was first fitted, as one of the most common ways of analysing time-to-event data (Cox, 1972). The model was fitted to understand the effect of the four binary

ECG characteristics on the composite, death, and constriction events using the HR as the measure of effect. The limitation of the standard Cox model is that it only uses data at baseline, which means that the binary ECG time-varying variables were only added at baseline, which is a problem because all the ECG data from other time points are discarded.

The overall fit of the Cox model was assessed by plotting the Cox-Snell residuals to provide more insight into how well the Cox model fits the data. The graphs of Cox-Snell residuals of all four binary ECG characteristics indicated that the standard Cox model was not a good fit for the data, which also indicated a limitation of the standard Cox model. Wide confidence intervals in the results of the Cox model were observed, which suggest little knowledge about the effect of four binary time-varying covariates on composite, death, and constriction events, and perhaps indicating that further investigation is needed. To further the investigation, an extended Cox model, also known as the Anderson and Gill model in the literature (Anderson & Gill, 1982), was fitted. Although the extended Cox model is ideal for fitting exogenous time-varying covariates, it is not appropriate when the interest lies in the strength of the association between the endogenous time-varying covariates and survival times, which was the case in this study.

The recommended joint model for longitudinal and time-to-event data was therefore fitted, which linked the survival model with the appropriate measurement model for the time-varying covariates. This modelling framework accounts for the limitations of the extended Cox model (Rizopoulos, 2012). Several approaches are taken towards jointly modelling time-to-event and longitudinal outcomes, with the initial one being a two-stage joint model approach. The model is known for its simplicity, and it was developed in two main steps. The GLMM was first fitted, then subject-specific predictions were extracted from the fitted GLMM and these were used as time-dependent covariates in the extended survival model. The GLMM was fitted using the LME4 Package in R, while the extended Cox model was implemented using the Survival Package.

The results were presented in two sub-models, namely the longitudinal and survival sub-models. It was found that the Cox-Snell residual plots of ECGrate, ECGqrsd, ECGpri, and ECGqtc (displayed in Figures 6.25 to 6.28) deviated from the reference line, which meant that the two-stage joint model was not a good fit for the IMPI data in either the composite or death event outcome. However, the Cox-Snell residual plots suggested otherwise in the constriction event outcome. This made sense as the two-stage model is known to produce biased results

because errors produced from the first step are not carried through to the next step of the extended Cox model (Dafni & Tsiatis, 1998; Sweeting & Thompson, 2012; Tsiatis & Davidian, 2001). Mauff et al. (2020) proposed a correction method using importance sampling weights for estimates derived from a two-stage approach, but this was deemed beyond the scope of this study.

The shared parameter joint model is another approach to joint modelling that has been an active research area. The four univariate shared parameter joint models were fitted for time-to-composite, time-to-death, and time-to-constriction events. In particular, the univariate shared parameter joint models were fitted for four binary time-varying covariates, i.e., ECGrate, ECGpri, ECGqrsd, and ECGqtc, to allow for comparison with other joint models. Additionally, the researcher wanted to investigate the association between ECG characteristics and study outcomes, such as death, constriction, or composite separately using a joint model framework. The univariate analysis presented in Tables 6.11 to 6.14 suggested a weak association between ECGrate and the risk of composite, death, and constriction event outcomes. However, there was a strong association between ECGpri and the risk of death and constriction survival outcomes, but there was no association in the composite event. Furthermore, there was no association between ECGqrsd and the risk of both composite and death events; however, there was an association with constriction. Finally, there was no association between ECGqtc and the risk of both composite and death events; however, there was an association between ECGqtc and constriction.

The analysis was extended to a multivariate shared parameter joint model. The model had four binary ECG longitudinal outcomes, which were modelled under the binomial assumption using GLMM. Parameter estimation was based on a Bayesian framework using MCMC, and convergency estimates were achieved. The model was fitted precisely to investigate if there was an association between the ECG characteristics and study outcomes. This was done by using association parameters for ECGs, which quantifies how the longitudinal ECG measurements are related to the risk of death, constriction, and composite events. The analysis of the ECG measurements revealed a notable slight increase in posterior means estimates for all three events. It further revealed that the time and prednisolone effects were statistically significant (see Table 6.15) in both the death and composite events. The main treatment, prednisolone, did not have a significant time interaction effect. In other words, prednisolone had no effect on the subjects with regard to time (see Table 6.15).

No association was observed between the ECGrate and the risk of composite, death, and constriction event outcomes, as shown by overlapping 95% confidence intervals. There was also no association between the ECGpri and the risk of composite, death, and constriction event outcomes. There was no association between the ECGqrsd and the risk of composite, death, and constriction event outcomes. Finally, there was no association between ECGqtc and the risk of composite and constriction event outcomes. However, there was an association between ECGqtc and the risk of death. It is worth noting that there was a marked difference between the ECG HR and prednisolone, which could have resulted from differences in ECG measurements taken or the effect of confounding variables that were fitted in the multivariate joint model, or even the treatment effect.

The study discovered that the model-checking technique exists in the frequentist framework when estimation was done using the likelihood ratio in the univariate and multivariate shared parameter joint model settings. However, when estimation of parameters was done using the Bayesian approach, which was the case in the shared parameter joint models and in this study, no model-checking tools are currently developed. The researcher relied on the assumption that the shared parameter model was a better fit for the IMPI trial data.

7.3 LIMITATIONS AND FUTURE RESEARCH

The study had several limitations. Firstly, while multivariate joint modelling for longitudinal and time-to-event data provided valuable insights into cardiovascular risk forecast and its associations using the IMPI trial data, multiple time-varying covariates increase convergence issues, requiring careful model selection and tuning. Secondly, doing simulation studies to compare and validate with the actual data (IMPI) outcome. In contrast, the delimitation of the study was that the analysis was restricted to patients with only ECG data and focused on patients with at least two ECG measurements. In addition, the ECG measurements from the IMPI trial dataset were transformed from continuous to categorical variables to make the variables clinically meaningful

A further area of research would be to conduct a sensitivity analysis using ECGs from patients without co-morbidity or TB and HIV. In addition, it would be of interest to examine other factors that may lead to death, such as heart pressure and viral load, in combination with the ECGs.

As mentioned in Section 6.4.4, there are currently no tools for model checking for the shared parameter joint model framework when estimation is done using a Bayesian approach. This is a potential area for further research. Mchunu et al. (2022) investigated an alternative association structure as opposed to the current value used in a joint model framework. However, there is little to no literature on the investigation of parametrisations in the joint multivariate model setting. This could also be an area of further research.

This study analysed IMPI trial data by means of the standard Cox model, extended Cox model, univariate joint model, and, the multivariate joint model setting taking into account the composite event of death, cardiac tamponade, or constriction as a single event of interest, and death, and constriction individually as the event of interest. However, there is a need to examine the contribution of competing risk analyses to the multivariate joint model. Finally, it is common in clinical trials to have data missingness at some point due to individual withdrawal from the study before completion. As mentioned in Section 2.4, the study assumed that some IMPI trial data were missing at random. One of the advantages of a joint modelling framework is that it addresses missingness. However, there is a need to extend the analysis into multivariate joint modelling to account for missing data in the longitudinal outcomes for further research.

BIBLIOGRAPHY

Adedinsewo, D., Carter, R. E., Attia, Z., Johnson, P., Kashou, A. H., Dugan, J. L., Albus, M., Sheele, J. M., Bellolio, F., Friedman, P. A., Lopez-Jimenez, F., & Noseworthy, P. A. (2020). Artificial intelligence-enabled ECG algorithm to identify patients with left ventricular systolic dysfunction presenting to the emergency department with dyspnea. *Circulation: Arrhythmia and Electrophysiology*, *13*(8), e008437.

Albert, P. S. (2000). A transitional model for longitudinal binary data subject to nonignorable missing data. *Biometrics*, *56*(2), 602–608.

Alvares, D., & Rubio, F. J. (2021). A tractable Bayesian joint model for longitudinal and survival data. *Statistics in Medicine*, *40*(19), 4213-4229.

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*, *10*(4), 1100–1120.

Andrinopoulou, E. R., Rizopoulos, D., Takkenberg, J. J., & Lesaffre, E. (2014). Joint modeling of two longitudinal outcomes and competing risk data. *Statistics in Medicine*, *33*(18), 3167–3178.

Andrinopoulou, E. R., Rizopoulos, D., Takkenberg, J. J., & Lesaffre, E. (2017). Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Statistical Methods in Medical Research*, *26*(4), 1787–1801.

Arjas, E. (1988). A graphical method for assessing goodness of fit in Cox's proportional hazards model. *Journal of the American Statistical Association*, *83*(401), 204–212.

Bang, C. N., & Okin, P. M. (2016). Does electrocardiography still have a place in hypertension? *Journal of Hypertension*, *34*(5), 842–844.

Barrett, J., & Su, L. (2017). Dynamic predictions using flexible joint models of longitudinal and time-to-event data. *Statistics in Medicine*, *36*(9), 1447–1460.

Bogun, F., Anh, D., Kalahasty, G., Wissner, E., Serhal, C. B., Bazzi, R., Weaver, W. D., & Schuger, C. (2004). Misdiagnosis of atrial fibrillation and its clinical consequences. *The American Journal of Medicine*, *117*(9), 636–642.

Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, *88*(421), 9–25.

Brilleman, S. L., Crowther, M. J., Moreno-Betancur, M., Buros Novik, J., Dunyak, J., Al-Huniti, N., ... & Wolfe, R. (2019). Joint longitudinal and time-to-event models for

multilevel hierarchical data. *Statistical Methods in Medical Research*, *28*(12), 3502-3515.

Brown, E., & Ibrahim, J. (2003). A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. *Biometrics*, *59*(2), 221–228.

Bycott, P., & Taylor, J. (1998). A comparison of smoothing techniques for CD4 data measured with error in a time-dependent cox proportional hazards model. *Statistics in Medicine*, *17*(18), 2061–2077.

Chi, Y. Y., & Ibrahim, J. G. (2006). Joint models for multivariate longitudinal and multivariate survival data. *Biometrics*, *62*(2), 432–445.

Chiou, S. H., Kang, S., & Yan, J. (2014). Fitting accelerated failure time models in routine survival analysis with R Package Aftgee. *Journal of Statistical Software*, *61*, 1–23.

Chishala, C. (2016). *The prevalence, determinants, natural history and impact of atrial fibrillation and atrial flutter in patients with tuberculosis pericarditis-insights from the IMPI trial* [Master's thesis, University of Cape Town].

Collett, D. (2003). *Modelling survival data in medical research*. Chapman Hall/CRC Texts in Statistical Science.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–202.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, *62*(2), 269–276.

Cox, D. R. (1979). A note on the graphical analysis of survival data. *Biometrika*, *66*(1), 188–190.

Dafni, U. G., & Tsiatis, A. A. (1998). Evaluating surrogate markers of clinical outcome when measured with error. *Biometrics*, 1445–1462.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Elashoff, R. M., Huang, X., & Li, G. (2010). A joint model of longitudinal and competing risks survival data with heterogeneous random effects and outlying longitudinal measurements. *Statistics and Its Interface*, *3*(2), 185–195.

Elashoff, R. M., Li, G., & Li, N. (2008). A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics*, *64*(3), 762–771.

Engel, B., & Keen, A. (1994). A simple approach for the analysis of generalized linear mixed models. *Statistica Neerlandica*, *48*(1), 1–22.

Fahrmeir, L., Tutz, G., Hennevogl, W., & Salem, E. (1994). *Multivariate statistical modelling based on generalized linear models* (vol. 425). Springer.

Faucett, C. L., Schenker, N., & Elashoff, R. M. (1998). Analysis of censored survival data with intermittently observed time-dependent binary covariates. *Journal of the American Statistical Association*, *93*(442), 427–437.

Faucett, C. L., & Thomas, D. C. (1996). Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. *Statistics in Medicine*, *15*(15), 1663–1685.

Ferro, E. G., Abrahams-Gessel, S., Jardim, T. V., Wagner, R., Gomez-Olive, F. X., Wade, A. N., Peters, F., Tollman, S., & Gaziano, T. A. (2021). Echocardiographic and electrocardiographic abnormalities among elderly adults with cardiovascular disease in rural South Africa. *Circulation: Cardiovascular Quality and Outcomes*, *14*(11), e007847.

Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, *94*(446), 496–509.

Follmann, D., & Wu, M. (1995). An approximate generalized linear model with random effects for informative missing data. *Biometrics*, 151–168.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Ha, I. D., Christian, N. J., Jeong, J.-H., Park, J., & Lee, Y. (2016). Analysis of clustered competing risks data using subdistribution hazard models with multivariate frailties. *Statistical Methods in Medical Research*, *25*(6), 2488–2505.

Harville, D. (1977). The use of linear-model methodology to rate high school or college football teams. *Journal of the American Statistical Association*, *72*(358), 278–289.

Hazra, A. (2017). Using the confidence interval confidently. *Journal of Thoracic Disease*, *9*(10), 4125.

Henderson, R., Diggle, P., & Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics*, *1*(4), 465–480.

Hickey, G. L., Philipson, P., Jorgensen, A., and Kolamunnage-Dona, R. (2016). Joint modelling of time-to-event and multivariate longitudinal outcomes: Recent developments and issues. *BMC Medical Research Methodology*, *16*(1), 1–15.

Hiller, L., Marshall, A., & Dunn, J. (2015). Assessing violations of the proportional hazards assumption in Cox regression: Does the chosen method matter? *Trials*, *16*(2), 134.

Hobert, J. P., & Casella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, *91*(436), 1461–1473.

Hogan, J. W., & Laird, N. M. (1998). Increasing efficiency from censored survival data by using random effects to model longitudinal covariates. *Statistical Methods in Medical Research*, *7*(1), 28–48.

Houben, R. M., & Dodd, P. J. (2016). The global burden of latent tuberculosis infection: A re-estimation using mathematical modelling. *PLoS Medicine*, *13*(10), e1002152.

Hsieh, F., Tseng, Y.-K., & Wang, J.-L. (2006). Joint modeling of survival and longitudinal data: Likelihood approach revisited. *Biometrics*, *62*(4), 1037–1043.

Huang, X., Li, G., Elashoff, R. M., & Pan, J. (2011). A general joint model for longitudinal measurements and competing risks survival data with heterogeneous random effects. *Lifetime Data Analysis*, *17*(1), 80–100.

Hunter, D. R., & Lange, K. (2004). A tutorial on MM algorithms. *The American Statistician*, *58*(1), 30–37.

Istolahti, T., Lyytikäinen, L.-P., Huhtala, H., Nieminen, T., Kähönen, M., Lehtimäki, T., Eskola, M., Anttila, I., Jula, A., Rissanen, H., Nikus, J., & Hernesniemi, J. (2021). The prognostic significance of T-wave inversion according to ECG lead group during long-term follow-up in the general population. *Annals of Noninvasive Electrocardiology*, *26*(1), e12799.

Kalbfleisch, J., & Prentice, R. (2002). *The statistical analysis of failure time data*. John Wiley & Sons.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53*(282), 457–481.

Karim, M. R., & Zeger, S. L. (1992). Generalized linear models with random effects: Salamander mating revisited. *Biometrics*, 631–644.

Król, A., Ferrer, L., Pignon, J.-P., Proust-Lima, C., Ducreux, M., Bouché, O., Michiels, S., & Rondeau, V. (2016). Joint model for left-censored longitudinal data, recurrent events and terminal event: Predictive abilities of tumor burden for cancer evolution with application to the FFCD 2000–05 trial. *Biometrics*, *72*(3), 907–916.

Lai, Y., Hayashida, M., & Akutsu, T. (2013). Survival analysis by penalized regression and matrix factorization. *The Scientific World Journal*, *23*, 632030.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 963–974.

Larsen, K. (2004). Joint analysis of time-to-event and multiple binary indicators of latent classes. *Biometrics*, *60*(1), 85–92.

Lawrence Gould, A., Boye, M. E., Crowther, M. J., Ibrahim, J. G., Quartey, G., Micallef, S., & Bois, F. Y. (2015). Joint modeling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modeling working group. *Statistics in medicine*, *34*(14), 2181-2195.

Lee, Y., Nelder, J., & Pawitan, Y. (2006). *Generalized linear models with random effects: Unified analysis via H-likelihood*. Routledge.

Lehtonen, A. O., Puukka, P., Varis, J., Porthan, K., Tikkanen, J. T., Nieminen, M. S., Huikuri, H. V., Anttila, I., Nikus, K., Kähönen, M., Jula, A., & Niiranen, T. (2016). Prevalence and prognosis of ECG abnormalities in normotensive and hypertensive individuals. *Journal of Hypertension*, *34*(5), 959–966.

Lemeshow, S., May, S., & Hosmer Jr, D. W. (2011). *Applied survival analysis: Regression modeling of time-to-event data*. John Wiley & Sons.

Li, N., Elashoff, R. M., & Li, G. (2009). Robust joint modeling of longitudinal measurements and competing risks failure time data. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, *51*(1), 19–30.

Li, N., Elashoff, R. M., Li, G., & Tseng, C.-H. (2012). Joint analysis of bivariate longitudinal ordinal outcomes and competing risks survival times with nonparametric distributions for random effects. *Statistics in Medicine*, *31*(16), 1707–1721.

Lin, H., Turnbull, B. W., McCulloch, C. E., & Slate, E. H. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data: Application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association*, *97*(457), 53–65.

Lipsitz, S. R., Fitzmaurice, G. M., Ibrahim, J. G., Gelber, R., & Lipshultz, S. (2002). Parameter estimation in longitudinal studies with outcome-dependent follow-up. *Biometrics*, *58*(3), 621–630.

Masoudi, F. A., Magid, D. J., Vinson, D. R., Tricomi, A. J., Lyons, E. E., Crounse, L., Ho, P. M., Peterson, P. N., & Rumsfeld, J. S. (2006). Implications of the failure to identify

high-risk electrocardiogram findings for the quality of care of patients with acute myocardial infarction: Results of the emergency department quality in myocardial infarction (EDQMI) study. *Circulation*, *114*(15), 1565–1571.

Mathur, M. L. (2006). Potential utility of Mycobacterium w vaccine in control of tuberculosis. *Current Respiratory Medicine Reviews*, *2*(2), 183–188.

Mauff, K., Steyerberg, E., Kardys, I., Boersma, E., & Rizopoulos, D. (2020). Joint models with multiple longitudinal outcomes and a time-to-event outcome: A corrected two-stage approach. *Statistics and Computing*, *30*(4), 999–1014.

Mayosi, B. M., Burgess, L. J., & Doubell, A. F. (2005). Tuberculous pericarditis. *Circulation*, *112*(23), 3608–3616.

Mayosi, B. M., Ntsekhe, M., Bosch, J., Pandie, S., Jung, H., Gumedze, F., Pogue, J., Thabane, L., Smieja, M., Francis, V., Joldersma, L., Thomas, K. M., Thomas, B., Awotedu, A. A., Magula, N. P., Naidoo, D. P., Damasceno, A., Banda, A. C., Brown, B., Manga, P., … et al. (2014). Prednisolone and *Mycobacterium indicus pranii* in tuberculous pericarditis. *New England Journal of Medicine*, *371*(12), 1121–1130.

Mayosi, B. M., Wiysonge, C. S., Ntsekhe, M., Gumedze, F., Volmink, J. A., Maartens, G., Aje, A., Thomas, B. M., Thomas, K. M., Awotedu, A. A., Thembela, B., Mntla, P., Maritz, F., Blacket, K. N., Nkouonlack, D. C., Burch, V. C., Rebe, K., Parrish, A., Sliwa, K., Vezi, B. Z., … et al. (2008). Mortality in patients treated for tuberculous pericarditis in sub-Saharan Africa. *South African Medical Journal*, *98*(1), 36–40.

Mayosi, B. M., Wiysonge, C. S., Ntsekhe, M., Volmink, J. A., Gumedze, F., Maartens, G., Aje, A., Thomas, B. M., Thomas, K. M., Awotedu, A. A., Thembela, B., Mntla, P., Maritz, F., Blacket, K. N., Nkouonlack, D. C., Burch, V. C., Rebe, K., Parrish, A., Sliwa, K., Vezi, B. Z., …et al. (2006). Clinical characteristics and initial management of patients with tuberculous pericarditis in the HIV era: The Investigation of the Management of Pericarditis in Africa (IMPI Africa) registry. *BMC Infectious Diseases*, *6*, 1–9.

McCulloch, C. E. (1994). Maximum likelihood variance components estimation for binary data. *Journal of the American Statistical Association*, *89*(425), 330–335.

McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association*, *92*(437):162–170.

McCulloch, C. E. (2003). *Generalized linear mixed models*. IMS.

McCulloch, C. E., & Searle, S. R. (2001). *Generalized, linear, and mixed models*. John Wiley & Sons.

McCullough, S. A., Goyal, P., Krishnan, U., Choi, J. J., Safford, M. M., & Okin, P. M. (2020). Electrocardiographic findings in Coronavirus Disease-19: Insights on mortality and underlying myocardial processes. *Journal of Cardiac Failure*, *26*(7), 626–632.

Mchunu, N. N., Mwambi, H. G., Rizopoulos, D., Reddy, T., & Yende-Zuma, N. (2022). Using joint models to study the association between CD4 count and the risk of death in TB/HIV data. *BMC Medical Research Methodology*, *22*(1), 295.

Mentel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*, *50*, 163–170.

Mieno, M. N., Yamaguchi, T., & Ohashi, Y. (2011). Alternative statistical methods for estimating efficacy of interferon beta-1b for multiple sclerosis clinical trials. *BMC Medical Research Methodology*, *11*(1), 1–6.

Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data*. Springer-Verlag.

Moore, D. F. (2016). *Applied survival analysis using R*. Springer.

Nelson, W. (1969). Hazard plotting for incomplete failure data. *Journal of Quality Technology*, *1*(1), 27–52.

Ntsekhe, M., Wiysonge, C. S., Gumedze, F., Maartens, G., Commerford, P. J., Volmink, J. A., & Mayosi, B. M. (2008). HIV infection is associated with a lower incidence of constriction in presumed tuberculous pericarditis: A prospective observational study. *PLoS One*, *3*(6), e2253.

Ntsekhe, M., Wiysonge, C., Volmink, J., Commerford, P., & Mayosi, B. (2003). Adjuvant corticosteroids for tuberculous pericarditis: Promising, but not proven. *QJM: An International Journal of Medicine*, *96*(8), 593–599.

Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, *58*(3), 545–554.

Prentice, R. L. (1982). Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*, *69*(2), 331–342.

Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. T., & Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, 541–554.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing* (3rd ed.). Cambridge University Press.

Prinja S., & Gupta N, V. R. (2010). Censoring in clinical trials: Review of survival analysis techniques. *Indian Journal of Community Medicine*, *35*(2), 217–217.

Proust-Lima, C., & Taylor, J. M. (2009). Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: A joint modeling approach. *Biostatistics*, *10*(3), 535–549.

Pulkstenis, E. P., Ten Have, T. R., & Landis, J. R. (1998). Model for the analysis of binary longitudinal pain data subject to informative dropout through remedication. *Journal of the American Statistical Association*, *93*(442), 438–450.

Reuter, H., Burgess, L., & Doubell, A. (2005). Epidemiology of pericardial effusions at a large academic hospital in South Africa. *Epidemiology & Infection*, *133*(3), 393–399.

Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., & Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngology—Head and Neck Surgery*, *143*(3), 331–336.

Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, *35*, 1–33.

Rizopoulos, D. (2012). *Joint models for longitudinal and time-to-event data: With applications in R*. CRC Press.

Rizopoulos, D., Verbeke, G., Lesaffre, E., & Vanrenterghem, Y. (2008). A two-part joint model for the analysis of survival and longitudinal binary data with excess zeros. *Biometrics*, *64*(2), 611–619.

Rue, M., Andrinopoulou, E. R., Alvares, D., Armero, C., Forte, A., & Blanch, L. (2017). Bayesian joint modeling of bivariate longitudinal and competing risks data: an application to study patient-ventilator asynchronies in critical care patients. *Biometrical Journal*, *59*(6), 1184-1203.

Sarkar, K., Chowdhury, R., & Dasgupta, A. (2017). Analysis of survival data: Challenges and algorithm-based model selection. *Journal of Clinical and Diagnostic Research*, *11*(6), LC14.

Sartor, C. I., Strawderman, M. H., Lin, X. H., Kish, K. E., McLaughlin, P. W., & Sandler, H. M. (1997). Rate of PSA rise predicts metastatic versus local recurrence after definitive radiotherapy. *International Journal of Radiation Oncology, Biology, Physics*, *38*(5), 941–947.

Schall, R. (1991). Estimation in generalized linear models with random effects. *Biometrika*, *78*(4), 719–727.

Searle, S. R. (1995). An overview of variance component estimation. *Metrika*, *42*(1), 215–230.

Self, S., & Pawitan, Y. (1992). Modeling a marker of disease progression and onset of disease. In N. P. Jewell, K. Dietz, & V. T. Farewell (Eds.). *AIDS epidemiology* (pp. 231–255). Springer.

Self, S. G., & Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, *82*(398), 605–610.

Stram, D., & Lee, J. (1995). Variance component testing in the longitudinal mixed effects model. *Biometrics*, *50*(4), 1171–1177.

Sweeting, M., & Thompson, S. (2012). Making predictions from complex longitudinal data, with application to planning monitoring intervals in a national screening programme. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *175*(2), 569–586.

Therneau, T. M., & Grambsch, P. M. (2000). The Cox model. In T. M. Therneau & P. M. Grambsch (Eds.). *Modeling survival data: Extending the Cox model* (pp. 39–77). Springer.

Therneau, T. M., Grambsch, P. M., & Fleming, T. R. (1990). Martingale-based residuals for survival models. *Biometrika*, *77*(1), 147–160.

Thijs, H., Molenberghs, G., & Verbeke, G. (2000). The milk protein trial: Influence analysis of the dropout process. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, *42*(5), 617–646.

Tsiatis, A. A., & Davidian, M. (2001). A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika*, *88*(2), 447–458.

Tsiatis, A. A., & Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica*, 14:809–834.

Tsiatis, A. A., Degruttola, V., & Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error: Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association*, *90*(429), 27–37.

Viljoen, C. A., Millar, R. S., Engel, M. E., Shelton, M., & Burch, V. (2017). Is computer-assisted instruction more effective than other educational methods in achieving ECG competence among medical students and residents? Protocol for a systematic review and meta-analysis. *BMJ Open*, *7*(12), e018811.

Viljoen, C. A., Millar, R. S., Hoevelmann, J., Muller, E., Hähnle, L., Manning, K., Naude, J., Sliwa, K., & Burch, V. C. (2021). Utility of mobile learning in electrocardiography. *European Heart Journal – Digital Health*, *2*(2), 202–214.

Vonesh, E., & Chinchilli, V. (1997). *Linear and non-linear models for the analysis of repeated measurements*. Marcel Decker.

Wang, Y., & Taylor, J. M. G. (2001). Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association*, *96*(455), 895–905.

Wiysonge, C. S., Ntsekhe, M., Thabane, L., Volmink, J., Majombozi, D., Gumedze, F., Pandie, S., & Mayosi, B. M. (2017). Interventions for treating tuberculous pericarditis. *Cochrane Database of Systematic Reviews*, *9*(9), CD000526.

Wolfinger, R., & O'Connell, M. (1993). Generalized linear mixed models: A pseudo-likelihood approach. *Journal of Statistical Computation and Simulation*, *48*(3-4), 233–243.

World Health Organization (WHO). (2022). *Global tuberculosis report 2021: Supplementary material*. https://iris.who.int/bitstream/handle/10665/360605/9789240046856-eng.pdf

Wulfsohn, M. S., & Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics*, 330–339.

Xu, J., & Zeger, S. L. (2001). Joint analysis of longitudinal data comprising repeated measures and times to events. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *50*(3), 375–387.

Xue, Y., & Schifano, E. D. (2017). Diagnostics for the Cox model. *Communications for Statistical Applications and Methods*, *24*(6), 583–604.

Yu, M., Law, N. J., Taylor, J. M., & Sandler, H. M. (2004). Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica*, *14*, 835–862.

APPENDICES

Appendix A: IMPI Dataset Continuous ECG Characteristics Analysis

Kaplan-Meier (KM) curves

The KM curve of the composite endpoint for the placebo versus prednisolone group is displayed in Figure A1. The placebo group exhibited the lowest survival rate compared to the prednisolone group. Furthermore, the log-rank test suggested a significant difference between the survival of the two groups with $p \leq 0.5$.

Figure A2 shows the KM of the death endpoint of the placebo versus the prednisolone group. It is evident from the graph that the survival rate of the placebo versus the prednisolone group did not exhibit much of a difference. The log-rank test further cemented that the two survival curves were the same with $p \geq 0.5$.

In Figure A3, which shows the KM curve of the placebo versus prednisolone by constriction endpoint, the survival curve of the prednisolone group exhibits a relatively high survival rate compared to the placebo. There is enough evidence that the survival rates of the two groups are indeed not the same, with a very small p-value of ≤ 0.5

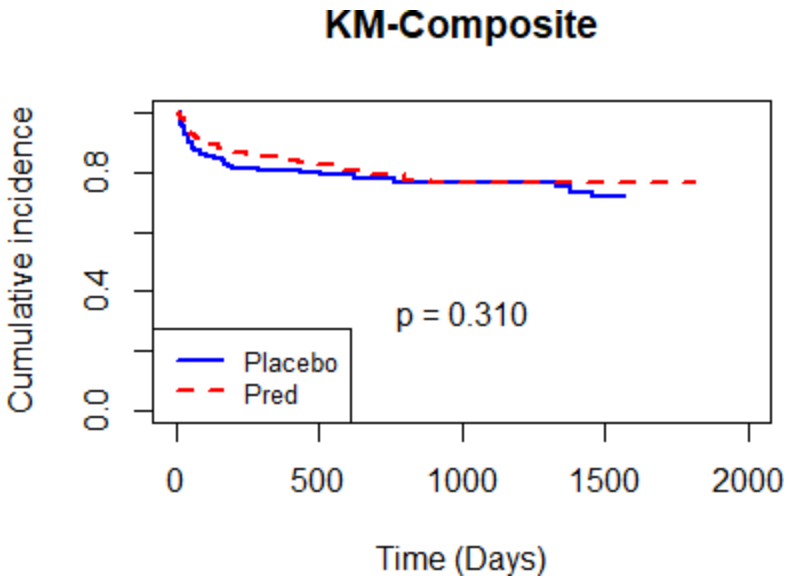

Figure A1: Illustration of the KM curves of placebo versus prednisolone group by composite endpoint

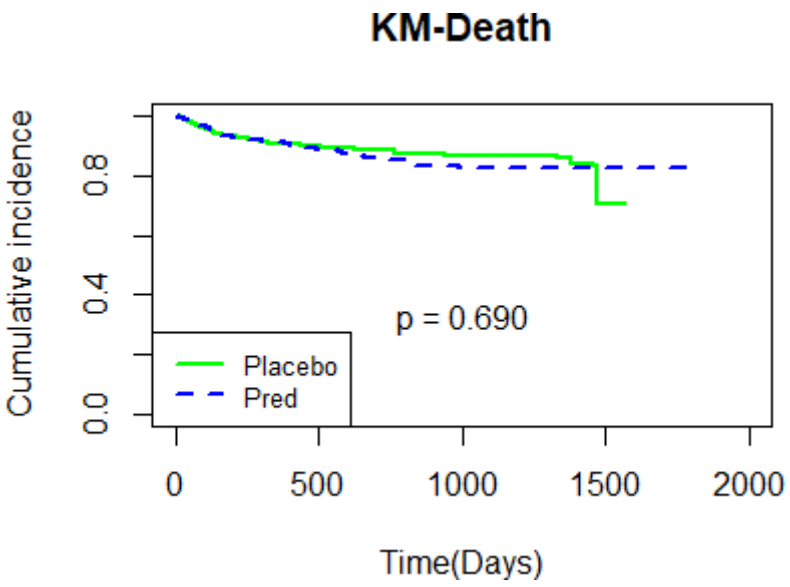

Figure A2: Illustration of the KM curves of placebo versus prednisolone group by death endpoint

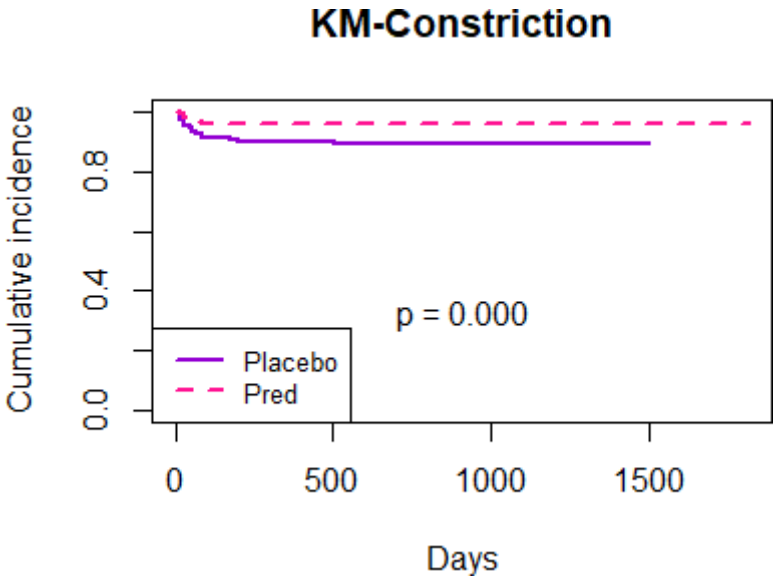

Figure A3: Illustration of the KM curves of placebo versus prednisolone group by constriction endpoint

Mean plot curves

The mean plots of the ECG characteristics are shown in Figures A4 to A7. The ECG measurements showed a linear trend and further decreased steadily with time from baseline until 11 months.

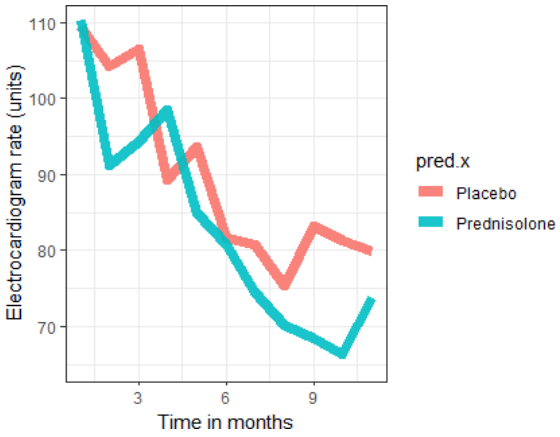

Figure A4: Mean plot of ECGrate over 11 months

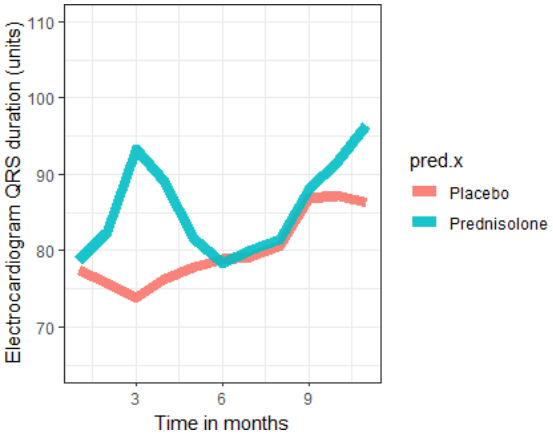

Figure A5: Mean plot of ECGqrsd over 11 months

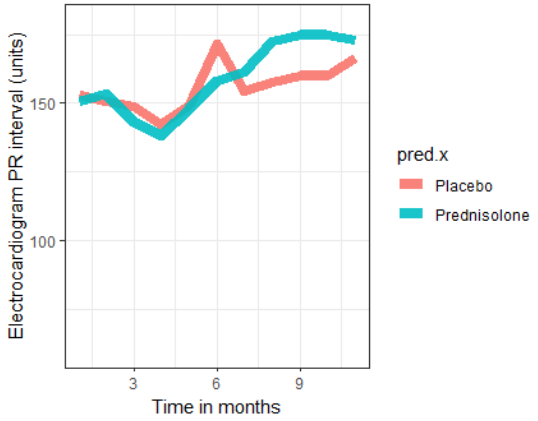

Figure A6: Mean plot of ECGpri over 11 months

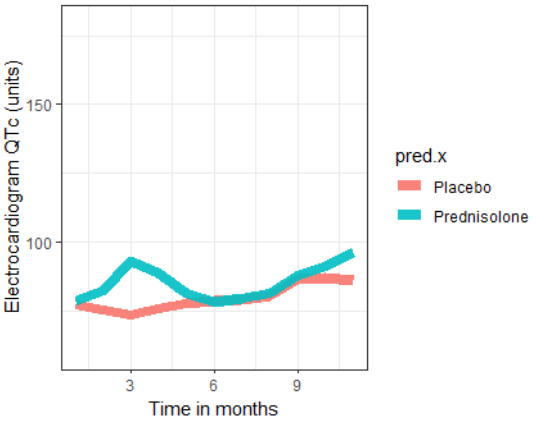

Figure A7: Mean plot of ECGqtc over 11 months

Analysis

The results of the extended Cox model for the time-to-composite, time-to-death, and time-to-constriction events are presented in Tables A1 to A3. Furthermore, the joint modelling results are presented in Tables A4 to A9.

Table A1: Extended Cox model for time-to-composite event

Variable	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value	Time to composite HR (95% CI)	p-value
ECGrate	0.991 (1.002 - 1.016)	0.014	-	-	-	-	-	-
ECGpri	-	-	1.005 (0.990 - 1.000)	0.082	-	-	-	-
ECGqtc	-	-	-	-	0.998 (0.999 - 1.005)	0.231	-	-
ECGqrsd	-	-	-	-	-	-	1.007 (0.980 - 1.007)	0.325
Prednisolone	1.244 (0.604 - 1.070)	0.135	1.233 (0.609 - 1.079)	0.150	1.236 (0.608 - 1.076)	0.146	1.225 (0.613 - 1.086)	0.164
Pericardiocentesis at randomisation (yes)	1.286 (0.532 - 1.137)	0.194	1.3019 (0.525 - 1.125)	0.175	1.259 (0.542 - 1.165)	0.239	1.272 (0.537 - 1.150)	0.215
Duration of symptoms (days)	0.995 (1.001 - 1.001)	0.006	0.995 (1.001 - 1.001)	0.012	0.995 (1.001 - 1.009)	0.010	0.995 (1.001 - 1.009)	0.009
Age (years)	0.996 (0.992 - 1.016)	0.498	0.997 (0.992 - 1.015)	0.576	0.997 (0.991 - 1.015)	0.626	0.997 (0.992 - 1.015)	0.572
NYHA class at study entry:								
II	0.766 (0.821 - 2.075)	0.260	0.760 (0.830 - 2.091)	0.245	0.749 (0.840 - 2.120)	0.222	0.750 (0.839 - 2.119)	0.224
III	0.555 (1.058 - 3.069)	0.030	0.521 (1.129 - 3.261)	0.016	0.538 (1.093 - 3.157)	0.022	0.537 (1.094 - 3.163)	0.022
IV	0.396 (1.343 - 4.754)	0.004	0.400 (1.328 - 4.720)	0.005	0.391 (1.358 - 4.806)	0.004	0.391 (1.361 - 4.811)	0.004
Creatinine (≥105)	0.590 (0.879 - 3.263)	0.115	0.616 (0.844 - 3.124)	0.147	0.609 (0.852 - 3.162)	0.139	0.601 (0.865 - 3.199)	0.128
HIV status (HIV+)	1.348 (0.528 - 1.043)	0.086	1.314 (0.542 - 1.069)	0.116	1.334 (0.533 - 1.054)	0.098	1.330 (0.535 - 1.057)	0.101
Weight (kg)	1.048 (0.841 - 1.081)	0.462	1.059 (0.832 - 1.071)	0.371	1.064 (0.829 - 1.065)	0.329	1.060 (0.832 - 1.070)	0.365
Palpable pulsus paradoxus (yes)	1.001 (0.675 - 1.475)	0.993	0.990 (0.682 - 1.494)	0.961	1.00 (0.676 - 1.478)	0.998	0.992 (0.682 - 1.490)	0.969
Systolic blood pressure (≥90)	1.694 (0.371 - 0.939)	0.026	1.644 (0.382 - 0.968)	0.036	1.696 (0.370 - 0.938)	0.026	1.673 (0.376 - 0.950)	0.029
Heart rate (≥100)	11.723 (0.011 - 0.675)	0.020	12.857 (0.010 - 0.622)	0.016	11.168 (0.011 - 0.708)	0.016	11.170 (0.011 - 0.708)	0.022
Effusion size:								
Medium (1-2 cm)	0.700 (0.551 - 3.700)	0.464	0.725 (0.530 - 3.588)	0.510	0.667 (0.574 - 3.913)	0.408	0.702 (0.548 - 3.703)	0.467
Small (<1 cm)	1.542 (0.213 - 1.975)	0.446	1.597 (0.205 - 1.911)	0.411	1.520 (0.216 - 2.007)	0.462	1.540 (0.213 - 1.981)	0.448
Chest X-ray pulmonary infiltrate (yes)	0940 (0.627 - 1.807)	0.819	0.954 (0.617 - 1.781)	0.862	0.975 (0.604 - 1.743)	0.924	0.968 (0.609 - 1.751)	0.905
Definite TB pericarditis status (yes)	0.705 (0.968 - 2.080)	0.073	0.658 (1.042 - 2.218)	0.030	0.660 (1.038 - 2.213)	0.031	0.664 (1.032 - 2.198)	0.034
Haemoglobin (≥10)	1.738 (0.076 - 4.343)	0.592	1.715 (0.077 - 4.425)	0.602	1.778 (0.074 - 4.252)	0.577	1.728 (0.076 - 4.380)	0.596
Peripheral oedema (yes)	0.688 (1.065 - 1.986)	0.018	0.662 (1.110 - 2.058)	0.009	0.664 (1.105 - 2.053)	0.010	0.664 (1.105 - 2.053)	0.010

Table A2: Extended Cox model for time-to-death event

Variable	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value	Time to death HR (95% CI)	p-value
ECGrate	0.993 (0.999 - 1.017)	0.131	-	-	-	-	-	-
ECGpri	-	-	1.007 (0.986 - 1.000)	0.049	-	-	-	-
ECGqtc	-	-	-	-	0.998 (0.997 - 1.006)	0.432	-	-
ECGqrsd	-	-	-	-	-	-	1.009 (0.974 - 1.008)	0.300
Prednisolone	0.987 (0.696 - 1.475)	0.947	0.978 (0.703 - 1.488)	0.908	0.982 (0.700 - 1.482)	0.923	0.977 (0.703 - 1.489)	0.905
Pericardiocentesis at randomisation (yes)	1.536 (0.402 - 1.056)	0.082	1.577 (0.389 - 1.033)	0.068	1.490 (0.413 - 1.091)	0.108	1.510 (0.408 - 1.074)	0.095
Duration of symptoms (days)	0.100 (0.994 - 1.006)	0.969	1.001 (0.994 - 1.001)	0.852	1.00 (0.994 - 1.006)	0.932	1.000 (0.994 - 1.006)	0.912
Age (years)	0.990 (0.994 - 1.028)	0.212	0.989 (0.994 - 1.028)	0.202	0.989 (0.994 - 1.028)	0.208	0.989 (0.994 - 1.028)	0.202
NYHA class at study entry:								
II	0.848 (0.636 - 2.189)	0.600	0.845 (0.639 - 2.191)	0.592	0.820 (0.658 - 2.264)	0.528	0.819 (0.658 - 2.267)	0.527
III	0.584 (0.861 - 3.408)	0.125	0.555 (0.908 - 3.577)	0.092	0.584 (0.860 - 3.413)	0.125	0.573 (0.876 - 3.478)	0.114
IV	0.332 (1.403 - 6.477)	0.005	0.337 (1.383 - 6.377)	0.005	0.330 (1.411 - 6.524)	0.004	0.315 (1.473 - 6.821)	0.003
Creatinine (≥105)	0.444 (1.336 - 3.802)	0.002	0.456 (1.300 - 3.692)	0.003	0.460 (1.295 - 3.657)	0.003	0.462 (1.289 - 3.637)	0.004
HIV status (HIV+)	0.905 (0.679 - 1.800)	0.687	0.891 (0.690 - 1.826)	0.642	0.898 (0.682 - 1.816)	0.668	0.910 (0.675 - 1.790)	0.706
Weight (kg)	1.150 (0.737 - 1.026)	0.098	1.138 (0.743 - 1.039)	0.129	1.156 (0.732 - 1.021)	0.087	1.148 (0.738 - 1.028)	0.103
Palpable pulsus paradoxus (yes)	1.118 (0.545 - 1.469)	0.659	1.140 (0.533 - 1.444)	0.607	1.122 (0.543 - 1.463)	0.650	1.137 (0.536 - 1.442)	0.611
Systolic blood pressure (≥90)	2.498 (0.228 - 0.702)	0.001	2.465 (0.231 - 0.714)	0.002	2.552 (0.223 - 0.687)	0.001	2.495 (0.229 - 0.702)	0.001
Heart rate (≥100)	1.120 (0.600 - 1.328)	0.576	1.046 (0.649 - 1.409)	0.821	1.068 (0.633 - 1.385)	0.742	1.067 (0.635 - 1.385)	0.745
Effusion size:								
Medium (1-2 cm)	0.884 (0.694 - 1.843)	0.621	0.90 (0.690 - 1.831)	0.640	0.884 (0.693 - 1.845)	0.622	0.871 (0.704 - 1.874)	0.578
Small (<1 cm)	1.667 (0.245 - 1.466)	0.262	1.730 (0.236 - 1.414)	0.230	1.704 (0.240 - 1.437)	0.244	1.649 (0.248 - 1.483)	0.273
Chest X-ray pulmonary infiltrate (yes)	1.038 (0.645 - 1.438)	0.854	1.021 (0.656 - 1.463)	0.920	1.038 (0.645 - 1.440)	0.856	1.016 (0.660 - 1.470)	0.940
Definite TB pericarditis status (yes)	0.855 (0.698 - 1.959)	0.552	0.819 (0.733 - 2.036)	0.443	0.813 (0.737 - 2.052)	0.428	0.822 (0.730 - 2.030)	0.451
Haemoglobin (≥10)	1.018 (0.646 - 1.493)	0.933	1.078 (0.611 - 1.406)	0.723	1.045 (0.632 - 1.450)	0.837	1.060 (0.622 - 1.430)	0.783
Peripheral oedema (yes)	0.687 (0.968 - 2.186)	0.071	0.673 (0.993 - 2.226)	0.054	0.663 (1.006 - 2.258)	0.047	0.678 (0.982 - 2.215)	0.061

Table A3: Extended Cox model for time-to-constriction event

Variable	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value	Time to constriction HR (95% CI)	p-value
ECGrate	0.976 (1.008 - 1.041)	0.004	-	-	-	-	-	-
ECGpri	-	-	1.016 (0.973 - 0.996)	0.008	-	-	-	-
ECGqtc	-	-	-	-	1.000 (0.994 - 1.005)	0.878	-	-
ECGqrsd	-	-	-	-	-	-	1.027 (0.944 - 1.005)	0.098
Prednisolone	2.647 (0.190 - 0.752)	0.006	2.290 (0.223 - 0.855)	0.016	2.348 (0.217 - 0.835)	0.013	2.244 (0.226 - 0.876)	0.019
Pericardiocentesis at randomisation (yes)	2.309 (0.197 - 0.951)	0.037	2.710 (0.166 - 0.819)	0.014	2.364 (0.192 - 0.931)	0.033	2.244 (0.203 - 0.979)	0.044
Duration of symptoms (days)	0.990 (1.003 - 1.017)	0.007	0.992 (1.000 - 1.015)	0.036	0.992 (1.001 - 1.016)	0.024	0.992 (1.001 - 1.015)	0.026
Age (years)	1.000 (0.975 - 1.025)	0.970	0.998 (0.977 - 1.028)	0.862	0.997 (0.978 - 1.028)	0.811	0.997 (0.978 - 1.028)	0.826
NYHA class at study entry:								
II	0.659 (0.531 - 4.331)	0.436	0.641 (0.545 - 4.475)	0.407	0.677 (0.523 - 4.176)	0.462	0.675 (0.522 - 4.207)	0.460
III	0.298 (1.062 - 10.615)	0.392	0.248 (1.299 - 12.509)	0.016	0.282 (1.142 - 10.994)	0.023	0.304 (1.057 - 10.211)	0.040
IV	0.416 (0.500 - 11.567)	0.274	0.432 (0.482 - 11.129)	0.295	0.416 (0.510 - 11.329)	0.268	0.443 (0.478 - 10.641)	0.304
Creatinine (≥105)	0.579 (0.726 - 4.106)	0.216	0.644 (0.659 - 3.659)	0.314	0.660 (0.644 - 3.567)	0.341	0.640 (0.666 - 3.665)	0.305
HIV status (HIV+)	2.750 (0.174 - 0.758)	0.007	2.364 (0.205 - 0.873)	0.020	2.527 (0.191 - 0.819)	0.013	2.424 (0.199 - 0.855)	0.017
Weight (kg)	0.900 (0.867 - 1.424)	0.407	0.943 (0.823 - 1.367)	0.650	0.944 (0.823 - 1.364)	0.652	0.918 (0.846 - 1.404)	0.504
Palpable pulsus paradoxus (yes)	0.962 (0.469 - 2.305)	0.923	0.861 (0.534 - 2.526)	0.705	0.925 (0.495 - 2.368)	0.842	0.920 (0.499 - 2.369)	0.834
Systolic blood pressure (≥90)	0.345 (0.381 - 22.034)	0.304	0.332 (0.398 - 22.876)	0.286	0.336 (0.394 - 22.508)	0.290	0.344 (0.385 - 21.946)	0.301
Heart rate (≥100)	1.007 (0.506 - 1.950)	0.985	0.796 (0.660 - 2.394)	0.487	0.787 (0.664 - 2.436)	0.469	0.831 (0.625 - 2.314)	0.580
Effusion size:								
Medium (1-2 cm)	0.873 (0.506 - 2.590)	0.745	0.957 (0.465 - 2.347)	0.915	0.902 (0.489 - 2.518)	0.805	0.904 (0.491 - 2.494)	0.808
Small (<1 cm)	1.763 (0.125 - 2.565)	0.461	2.192 (0.010 - 2.082)	0.311	2.003 (0.109 - 2.268)	0.368	1.859 (0.112 - 2.438)	0.421
Chest X-ray pulmonary infiltrate (yes)	0.834 (0.637 - 2.256)	0.573	0.773 (0.694 - 2.410)	0.419	0.729 (0.738 - 2.548)	0.318	0.779 (0.689 - 2.394)	0.431
Definite TB pericarditis status (yes)	0.404 (1.119 - 5.476)	0.025	0.325 (1.394 - 6.792)	0.05	0.338 (1.352 - 6.495)	0.007	0.364 (1.250 - 6.024)	0.012
Haemoglobin (≥10)	0.510 (0.953 - 4.038)	0.067	0.609 (0.812 - 3.317)	0.168	0.583 (0.841 - 3.471)	0.138	0.609 (0.808 - 3.340)	0.170
Peripheral oedema (yes)	0.823 (0.608 - 2.433)	0.580	0.823 (0.608 - 2.433)	0.375	0.823 (0.608 - 2.433)	0.291	0.725 (0.704 - 2.708)	0.349

Table A4: Longitudinal parameter estimates for ECGrate

Variable	Time to composite Estimate (SE)	t-value	p-value	Time to death Estimate (SE)	t-value	p-value	Time to constriction Estimate (SE)	t-value	p-value
Longitudinal sub-model									
Intercept	107.969 (0.903)	119.607	0.000	107.866 (0.898)	120.117	0.000	107.986 (0.903)	119.632	0.000
Prednisolone	-0.860 (1.241)	-0.693	0.489	-1.128 (1.238)	-0.911	0.363	-1.026 (1.241)	-0.827	0.409
Time	-0.467 (0.030)	-15.626	0.000	-0.398 (0.027)	-15.021	0.000	-0.454 (0.029)	-15.522	0.000
AIC	10260.49			10567.62			10308.01		
BIC	10295.82			10603.15			10343.37		

Table A5: Survival parameter estimates for ECGrate

Variable	Time to composite HR (95% CI)	SE	p-value	Time to death HR (95% CI)	SE	p-value	Time to constriction HR (95% CI)	SE	p-value
Survival sub-model									
Prednisolone	1.225 (0.613 - 1.086)	0.146	0.164	0.974 (0.706 - 1.495)	0.704	0.890	2.492 (0.202 - 0.796)	0.350	0.009
Pericardiocentesis at randomisation (yes)	1.284 (0.533 - 1.139)	0.194	0.198	1.526 (0.404 - 1.062)	0.906	0.086	2.313 (0.197 - 0.950)	0.402	0.037
Duration of symptoms (days)	0.995 (1.001 - 1.009)	0.002	0.007	1.000 (0.994 - 1.006)	0.151	0.984	0.990 (1.003 - 1.017)	0.004	0.008
Age (years)	0.996 (0.992 - 1.016)	0.006	0.522	0.989 (0.994 - 1.028)	0.420	0.208	1.001 (0.974 - 1.025)	0.013	0.951
NYHA class at study entry:	0.767 (0.820 - 2.071)			0.842 (0.640 - 2.203)			0.663 (0.529 - 4.302)		
II		0.236	0.262		1.161	0.587		0.535	0.442
III	0.547 (1.074 - 3111)	0.271	0.026	0.580 (0.867 - 3.435)	1.292	0.120	0.295 (1.075 - 10.713)	0.587	0.037
IV	0.385 (1.379 - 4.884)	0.323	0.003	0.325 (1.432 - 6.13)	1.435	0.004	0.383 (0.540 - 12.607)	0.804	0.233
Creatinine (≥105)	0.582 (0.891 - 3.313)	0.335	0.106	0.452 (1.315 - 3.728)	0.978	0.003	0.589 (0.715 - 4.035)	0.442	0.230
HIV status (HIV+)	1.334 (0.533 - 1.053)	0.173	0.096	0.898 (0.684 - 1.813)	0.915	0.666	2.681 (0.179 - 0.777)	0.374	0.008
Weight (kg)	1.052 (0.838 - 1.077)	0.064	0.421	1.152 (0.736 - 1.024)	1.140	0.094	0.904 (0.864 - 1.417)	0.126	0.424
Palpable pulsus paradoxus (yes)	0.993 (0.681 - 1.489)	0.110	0.972	1.107 (0.550 - 1.485)	0.932	0.689	0.955 (0.472 - 2.327)	0.407	0.909
Systolic blood pressure (≥90)	1.663 (0.378 - 0.957)	0.237	0.032	2.467 (0.231 - 0.711)	1.056	0.002	0.332 (0.396 - 22.918)	1.035	0.287
Heart rate (≥100)	11.202 (0.011 - 0.706)	1.055	0.593	1.095 (0.616 - 1.354)	0.739	0.650	1.018 (0.501 - 1.927)	0.344	0.959
Effusion size:	0.696 (0.554 - 3.725)			0.878 (0.690 - 1.856)			875 (0.505 - 2.587)		
Medium (1-2 cm)		0.486	0.456		0.916	0.600		0.417	0.749
Small (< 1 cm)	1.520 (0.216 - 2.002)	0.568	0.461	1.636 (0.250 - 1.494)	1.678	0.280	1.745 (0.127 - 2.591)	0.770	0.470
Chest X-ray pulmonary infiltrate (yes)	0.943 (0.625 - 1.710)	0.270	0.829	1.032 (0.649 - 1.446)	0.752	0.876	0.836 (0.637 - 2.250)	0.322	0.577
Definite TB pericarditis status (yes)	0.697 (0.980 - 2.102)	0.195	0.063	0.839 (0.713 - 1.993)	0.965	0.503	0.408 (1.110 - 5.422)	0.405	0.027
Haemoglobin (≥10)	1.735 (0.076 - 4.349)	1.031	0.593	1.021 (0.645 - 1.490)	0.787	0.925	0.497 (0.975 - 4.159)	0.370	0.058
Peripheral oedema (yes)	0.682 (1.074 - 1.910)	0.159	0.016	0.679 (0.981 - 2.211)	0.762	0.062	0.835 (0.598 - 2.397)	0.354	0.610

Table A6: Longitudinal parameter estimates for ECGpri

Variable	Time to composite Estimate (SE)	t-value	p-value	Time to death Estimate (SE)	t-value	p-value	Time to constriction Estimate (SE)	t-value	p-value
Longitudinal sub-model									
Intercept	5.008 (0.008)	621.307	0.000	5.008 (0.008)	620.184	0.000	5.008 (0.008)	619.150	0.000
Prednisolone	-0.011 (0.11)	-0.989	0.323	-0.009 (0.011)	-0.773	0.440	-0.010 (0.011)	-0.861	0.340
Time	0.001 (0.000)	4.214	0.000	0.000 (0.000)	3.213	0.002	0.000 (0.000)	4.133	0.000
AIC	-621.769			-639.5182			-620.0855		
BIC	-586.4364			-603.9817			-584.7226		

Table A7: Survival parameter estimates for ECGpri

Variable	Time to composite HR (95% CI)	SE	p-value	Time to death HR (95% CI)	SE	p-value	Time to constriction HR (95% CI)	SE	p-value
Survival sub-model	1.259 (0.597 - 1.058)			1.002 (0.685 - 1.453)			2.416 (0.777 - 0.811)		
Prednisolone		0.146	0.115			0.991		0.344	0.010
Pericardiocentesis at randomisation (yes)	1.313 (0.520 - 1.116)	0.195	0.162	1.597 (0.384 - 1.021)	0.250	0.061	2.840 (0.580 - 0.787)	0.410	0.012
Duration of symptoms (days)	0.995 (1.001 - 1.008)	0.002	0.013	1.001 (0.993 - 1.005)	0.003	0.847	0.992 (1.000 - 1.015)	0.004	0.037
Age (years)	0.997 (0.991 - 1.015)	0.006	0.596	0.990 (0.994 - 1.023)	0.008	0.214	0.998 (3.593 - 1.028)	0.013	0.876
NYHA class at study entry:	0.764 (0.823 - 2.078)			0.856 (0.631 - 2.162)			0.640 (2.005 - 4.477)		
II			0.255		0.314	0.622		0.537	
III	0.519 (1.134 - 3.270)	0.270	0.015	0.550 (0.916 - 3.606)		0.087	0.254 (1.273 - 12.202)	0.577	0.017
IV	0.403 (1.316 - 4.689)	0.324	0.005	0.341 (1.363 - 6.300)	0.391	0.006	0.442 (1.722 - 10.953)	0.804	0.309
Creatinine (≥105)	0.609 (0.853 - 3.158)	0.334	0.138	0.453 (1.307 - 3.720)	0.267	0.003	0.637 (2.449 - 3.708)	0.438	0.302
HIV status (HIV+)	1.310 (0.543 - 1.072)	0.173	0.120	0.881 (0.699 - 1.846)	0.248	0.608	2.336 (0.763 - 0.883)	0.369	0.022
Weight (kg)	1.057 (0.834 - 1.073)	0.064	0.386	1.135 (0.745 - 1.042)	0.086	0.140	0.942 (3.029 - 1.368)	0.130	0.646
Palpable pulsus paradoxus (yes)	0.985 (0.686 - 1.502)	0.200	0.941	1.141 (0.531 - 1.445)	0.255	0.604	0.828 (2.040 - 2.629)	0.397	0.635
Systolic blood pressure (≥90)	1.629 (0.386 - 0.978)	0.237	0.040	2.447 (0.232 - 0.719)	0.288	0.002	0.323 (1.501 - 23.462)	1.034	0.275
Heart rate (≥100)	10.918 (0.012 - 0.724)	1.055	0.023	1.039 (0.653 - 1.418)	0.198	0.846	0.804 (2.406 - 2.366)	0.328	0.506
Effusion size:	0.735 (0.522 - 3.539)			0.893 (0.687 - 1.823)			1.021 (1.598 - 2.209)		
Medium (1-2 cm)						0.651		0.415	0.960
Small (< 1 cm)	1.606 (0.204 - 1.900)		0.405	1.716 (0.238 - 1.424)		0.236	2.242 (1.323 - 2.036)	0.775	0.297
Chest X-ray pulmonary infiltrate (yes)	0.947 (0.621 - 1.795)	0.271	0.840	1.013 (0.661 - 1.475)	0.205	0.950	0.780 (2.528 - 2.391)	0.318	0.435
Definite TB pericarditis status (yes)	0.659 (1.040 - 2.213)	0.193	0.031	0.826 (0.726 - 2.018)	0.261	0.464	0.326 (1.390 - 6.787)	0.405	0.006
Haemoglobin (≥10)	1.680 (0.078 - 4.522)	1.034	0.616	1.074 (0.614 - 1.412)	0.212	0.737	0.611 (2.984 - 3.302)	0.358	0.169
Peripheral oedema (yes)	0.662 (1.109 - 2.056)		0.009	0.673 (0.992 - 2.227)	0.206	0.055	0.733 (2.549 - 2.690)	0.346	0.362

Table A8: Longitudinal parameter estimates for ECGqrsd

	Time to composite			Time to death			Time to constriction		
Variable	**Estimate (SE)**	**t-value**	**p-value**	**Estimate (SE)**	**t-value**	**p-value**	**Estimate (SE)**	**t-value**	**p-value**
Longitudinal sub-model									
Intercept	4.338 (0.006)	681.354	0.000	4.338 (0.006)	683.633	0.000	4.338 (0.006)	680.722	0.000
Prednisolone	-0.860 (0.009)	1.782	0.075	0.015 (0.009)	1.726	0.085	0.016 (0.009)	1.749	0.081
Time	-0.467 (0.000)	2.026	0.044	0.000 (0.000)	1.701	0.091	0.000 (0.000)	1.532	
AIC	-1174.564			-1201.021			-1172.915		
BIC	-1139.231			-1165.485			-1137.552		

Table A9: Survival parameter estimates for ECGqrsd

	Time to composite			Time to death			Time to constriction		
Variable	**HR (95% CI)**	**SE**	**p-value**	**HR (95% CI)**	**SE**	**p-value**	**HR (95% CI)**	**SE**	**p-value**
Survival sub-model	1.215 (0.617 - 1.096)	0.146		0.970 (0.708 - 1.502)			2.149 (0.865 - 0.921)		
Prednisolone			0.183		0.192	0.872		0.348	0.028
Pericardiocentesis at randomisation (yes)	1.274 (0.536 - 1.148)	0.194	0.212	1.512 (0.408 - 1.073)	0.247	0.094	2.262 (0.739 - 0.973)	0.402	0.043
Duration of symptoms (days)	0.995 (1.001 - 1.009)	0.002	0.009	1.000 (0.994 - 1.006)	0.003	0.907	0.992 (1.001 - 1.015)	0.004	0.026
Age (years)	0.997 (0.992 - 1.015)	0.006	0.571	0.989 (0.994 - 1.028)	0.008	0.200	0.997 (3.597 - 1.028)	0.013	0.835
NYHA class at study entry:	0.752 (0.837 - 2.114)	0.236		0.823 (0.655 - 2.256)			0.678 (1.914 - 4.188)		
II			0.228		0.316	0.537		0.532	0.464
III	0.539 (1.092 - 3.157)	0.271	0.022	0.575 (0.872 - 3.464)	0.352	0.116	0.307 (1.049 - 10.119)	0.578	0.041
IV	0.393 (1.355 - 4.791)	0.322	0.004	0.317 (1.466 - 6.795)	0.391	0.003	0.457 (1.704 - 10.348)	0.792	0.323
Creatinine (≥105)	0.599 (0.867 - 3.214)	0.334	0.125	0.462 (1.287 - 3.634)	0.265	0.004	0.639 (2.455 - 3.671)	0.435	0.303
HIV status (HIV+)	1.327 (0.536 - 1.059)	0.174	0.103	0.908 (0.676 - 1.794)	0.249	0.699	2.406 (0.736 - 0.863)	0.373	0.019
Weight (kg)	1.061 (0.832 - 1.069)	0.064	0.360	1.150 (0.736 - 1.026)	0.085	0.098	0.917 (3.113 - 1.404)	0.129	0.504
Palpable pulsus paradoxus (yes)	0.991 (0.682 - 1.493)	0.200	0.963	1.135 (0.537 - 1.446)	0.205	0.932	0.917 (1.836 - 2.381)	0.399	0.829
Systolic blood pressure (≥90)	1.673 (0.376 - 0.950)	0.236	0.030	2.497 (0.228 - 0.702)	0.286	0.001	0.345 (1.415 - 21.893)	1.031	0.301
Heart rate (>100)	11.166 (0.011 - 0.708)	1.055	0.023	1.061 (0.638 - 1.392)	0.199	0.767	0.835 (2.289 - 2.306)	0.334	0.589
Effusion size:	0.699 (0.550 - 3.716)	0.487		0.869 (0.705 - 1.877)			0.906 (1.801 - 2.489)		
Medium (1-2 cm)			0.463		0.250	0.575		0.415	0.415
Small (< 1 cm)	1.531 (0.214 - 1.993)	0.569	0.454	1.642 (0.249 - 1.490)	0.457	0.277	1.836 (0.442 - 2.469)	0.771	0.431
Chest X-ray pulmonary infiltrate (yes)	0.968 (0.609 - 1.752)	0.270	0.905	1.018 (0.658 - 1.467)	0.205	0.932	0.783 (2.519 - 2.382)	0.318	0.442
Definite TB pericarditis status (yes)	0.663 (1.034 - 2.203)	0.193	0.033	0.819 (0.732 - 2.037)	0.261	0.444	0.364 (1.251 - 6.020)	0.401	0.012
Haemoglobin (≥10)	1.734 (0.076 - 4.365)	1.033	0.594	1.058 (0.623 - 1.434)	0.213	0.791	0.609 (2.964 - 3.345)	0.363	0.172
Peripheral oedema (yes)	0.664 (1.103 - 2.053)	0.158	0.010	0.675 (0.987 - 2.227)	0.208	0.058	0.727 (2.577 - 2.700)	0.344	0.355

Appendix B: R Code

Multivariate shared parameter joint model

This section presents the R code for jointly modelling the four ECG characteristics' longitudinal binary outcomes (ECGrate, ECGpri, ECGqtc, and ECGqrsd) and the three events (composite, death, and constriction) using the shared parameter model approach. The same code was used for all other events.

```
#-------------------------------------------------------------------------
#-------------------JM Model(shared parameter Model)----------------
lmeFit <- mixed_model(ecgrate_binary ~ predx+wks
, data = lmm_ecg,random = ~wks|id1, na.action = na.exclude, family = binomial())
lmeFit1 <- mixed_model(ecgpri_binary ~ predx+wks
, data = lmm_ecg,random = ~1|id1, na.action = na.exclude, family= binomial())
lmeFit2 <- mixed_model(ecgqrsd_binary ~ predx+wks
, data = lmm_ecg,random = ~1|id1, na.action = na.exclude, family = binomial())
lmeFit3 <- mixed_model(ecgqtc_binary ~ predx+wks
, data = lmm_ecg,random = ~1|id1, na.action = na.omit, family = binomial())
CoxFit <- coxph(Surv(timeconstriction, eventconstriction) ~ predx+perf
+hivstatx+syblood
+definitetbp+iperipheraloedema+effusizex+ pulmonaryinfiltrates
+fpalpableparadoxus+agex+nyhaclassx, data = Surv_ecg)
jointFit <- jm(CoxFit, list(lmeFit, lmeFit1, lmeFit2,lmeFit3), time_var =
"wks", n_iter = 12000L, n_burnin = 2000L, n_thin = 5L)
summary(jointFit)
```