Sensitivity analysis approaches for incomplete longitudinal data in a multi-centre clinical trial

Abdul-Karim Iddrisu

Thesis presented for the degree of

DOCTOR OF PHILOSOPHY

in the Department of Statistical Sciences

University of Cape Town

Supervisor: Associate Professor Freedom N. Gumedze

December 2019

Declaration

I have read and understood the School'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 I have acknowledged all results and quotations from the published or unpublished work of other people.

Mr. Abdul-Karim Iddrisu

Signed by candidate

Signature:.....Date: 9th September 2019

Dedication

I dedicate this thesis to my wife Fuseina, daughter Zakiah and son Abdulai for their support and patience during the years of my research.

Abstract

The first major contribution of the thesis is the development of sensitivity analysis strategy for dealing with incomplete longitudinal data. The second important contribution is setting up of simulation experiment to evaluate the performance of some of the sensitivity analysis approaches. The third contribution is that the thesis offers recommendations on which sensitivity analysis strategy to use and in what circumstance. It is recommended that when drawing statistical inferences in the presence of missing data, methods of analysis based on plausible scientific assumptions should be used. One major issue is that such assumptions cannot be verified using the data at hand. In order to verify these assumptions, sensitivity analysis should be performed to investigate the robustness of statistical inferences to plausible alternative assumptions about the missing data. The thesis implemented various sensitivity analysis strategies to incomplete longitudinal CD4 count data in order to investigate the effect of tuberculosis pericarditis (TBP) treatment on CD4 count changes over time. The thesis achieved the first contribution by formulating primary analysis (which assume that the data are missing at random) and then conducting sensitivity analyses to assess whether statistical inferences under the primary analysis model are sensitive to models that assume that the data are not missing at random. The second contribution was achieved via simulation experiment involving formulating hypotheses on how sensitivity analysis strategies would performed under varying rate of missing values and model mis-specification (when the model is mis-specified). The third contribution was achieved based on our experience from the development and application of the sensitivity analysis strategies as well as the simulation experiment. Using the CD4 count data, we observed that statistical inferences under the primary analysis formulation are robust to the sensitivity analyses formulations, suggesting that the mechanism that generated the missing CD4 count measurements is likely to be missing at random. The results also revealed that TBP does not interact with the HIV/AIDS treatment and that TBP treatment had no significant effect on CD4 count changes over time. We have observed in our simulation results that the sensitivity analysis strategies produced unbiased statistical inferences except when a strategy is inappropriately applied in a given trial setting and also, when a strategy is mis-specified. Although the methods considered were applied to data in the IMPI trial setting, these methods can also be applied to clinical trials with similar settings. A sensitivity analysis strategy may not necessarily give bias results because it has been mis-specified, but it may also be that it has been applied in a wrongly defined trial setting. We therefore strongly encourage analysts to carefully study these sensitivity analysis frameworks together with a clearly and precise definition of the trial objective in order to decide on which sensitivity analysis strategy to use.

List of publications

Parts of this thesis have been published:

Paper 1 (Chapter 4): **Abdul-Karim Iddrisu and Freedom Gumedze (2018)**: Application of sensitivity analysis to incomplete longitudinal CD4 count data, Journal of Applied Statistics 46(4), 754-769, doi:https://doi.org/10.1080/02664763.2018.1510476.

Paper 2 (Chapter 5): **Abdul-Karim Iddrisu and Freedom Gumedze (2018):** An application of a pattern-mixture model with multiple imputation for the analysis of longitudinal trials with protocol deviations, BMC Medical Research Methodology 19(1), 1-10, doi: https://doi.org/10.1186/s12874-018-0639-y.

Paper 3 (Chapter 6): **Abdul-Karim Iddrisu and Freedom Gumedze (2019):** Sensitivity analysis for the generalized shared-parameter model, Journal of Biopharmaceutical Statistics, 1-19, doi:https://doi.org/10.1080/10543406.2019.1632875.

Acknowledgments

I would like to thank Associate Professor Freedom N. Gumedze for his excellent supervision skills as well as his enthusiastic guidance into the world of research.

Thank you to Associate Professor Francesca Little and Associate Professor Freedom N. Gumedze for giving me training in Biostatistics and introduction to longitudinal data and issues associated with such data, as well as methods of analysis.

I would like to thank the South African Centre for Epidemiological Modelling and Analysis (SACEMA) for funding the project, and The Academy of Medical Sciences and the National Research Foundation for partially funding this research. Thank you to the Mayosi Research Group, Department of Medicine, University of Cape Town, for providing the data for the study.

I would like to thank Prof. Jane Hutton, University of Warwick, Department of Statistics, UK, for their invaluable comments on the pattern-mixture model with multiple imputation (PM-MI) (Chapter 5). The author would also like to thank Suzie Cro of the London School of Hygiene and Tropical Medicine, UK, for making software available as well as offering valuable suggestions for the software's implementation under the PM-MI approach.

I would also like to thank Dr. Jennifer Rogers, Director of Statistical Consultancy Services, Department of Statistics, University of Oxford, UK, for taking the time from her busy schedule to proof-read the paper on sensitivity analyses for the selection model (Chapter 4).

I would like to thank the postgraduate students as well as the academic and administrative staff of the Department of Statistical Sciences, for their continuing support and encouragement throughout my PhD study period.

Abdul-Karim Iddrisu

December 2019

List of abbreviations

ACM: Approximate Conditional Model

ACMV: Available case missing values

ART: Anti-retroviral therapy

BOCF: Baseline observation carried forward

CC: Complete case analysis

CCMV: Complete case missing values

CHMP: Committee for medicinal products for Human Use

CIR: Copy increment from reference

CR: Copy reference

DK: Diggle and Kenward

EM: Expectation maximization

GEE: Generalized estimating equations

GLMM: Generalized linear mixed effect model

GSPM: Generalized shared-parameter model

IPW: Inverse probability weighting

J2R: Jump to reference

LD: Likelihood displacement

LOA: Last observation analysis

LOCF: Last observation carried forward

LMCF: Last mean carried forward

LMM: Linear mixed effect model

MAR: Missing at random

MCMC: Markov chain Monte Carlo

MCAR: Missing completely at random

MI: Multiple imputation

ML: Maximum likelihood

NCMV: Neighboring case missing values

NMAR: Not missing at random

NRC: National Research Council

PM-MI: Pattern-mixture model with multiple imputation

PMM: Pattern-mixture model

REML: Restricted maximum likelihood

RMSE: Mean square error

SI: Simple imputation

SeM: Selection model

SPM: Shared-parameter model

TBP: Tuberculosis pericarditis

wGEE: Weighted generalized estimating equations

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Longitudinal data typically occur in repeated measures designs, where measurements are recorded repeatedly on a response variable of interest for each patient at some selected scheduled visits. In addition to the response, covariates that are considered to be associated with the response may also be recorded. The focus of such designs is on how the response depends on the covariates over time. It is common for some of the measurements that were planned to be measured in longitudinal study designs to be missing. The occurrence of missing data in longitudinal studies is often higher than in cross-sectional studies (see Molenberghs and Fitzmaurice in Fitzmaurice et al., 2008, Ch. 17). This is because not all of the planned measurements are taken at all scheduled visits. In cross-sectional studies, data are collected from a group of subjects at a single point in time. For longitudinal study designs, missing data can occur at any scheduled visit. This means that the subject's response can be missing at one follow-up time and then be measured at a later follow-up time. Longitudinal studies also have missing data due to withdrawal/attrition. This results in different missingness patterns (Little, 1995; Rubin, 1996; Van Buuren, 2007, 2012).

It is known that patients may withdraw from longitudinal studies due to adverse effects of the treatment or protocol deviation (NRC, 2010; Carpenter et al., 2013; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Ayele et al., 2014). A protocol deviation is a study conduct in which the rules and procedure of conducting a clinical trial are altered by the investigator during the trial period. However, such rules and procedure can only be altered after an investigator has written and obtained authorization to do so. Depending on the study protocol, such deviations may be defined as (1) poor compliance with the intervention, or withdrawal from the intervention, (2) unblinding either of the intervention, or the evaluation and (3) moving to partial compliance with treatment or dropout such that no further information is recorded on the patient (Carpenter et al., 2013).

Missing data can affect precision of parameter estimates and can introduce bias. When longitudinal data are incomplete, there is loss of information and a reduction in the precision with which longitudinal change can be estimated (Molenberghs and Kenward, 2007; Fitzmaurice et al., 2008). This means that the reduction in precision is associated with the amount of missing data and to some extent, influenced by the method of analysis. However, the degree of bias that missing data may introduce into the analysis is of great concern. Bernhard et al. (1998) stated that missing data are a potential source of bias when analyzing clinical trial, and standard reporting requirements

include accounting for the missing data in the statistical analysis. This applies especially when the number of missing values are substantial. Kidson and Trenberth (1988); Molenberghs and Kenward (2007) also noted that missing data can introduce bias, thereby misleading statistical inferences about the changes in the response over time. This means that, to obtain valid statistical inferences, any method of analysis will be required to take into account the reasons for missingness (Little, 1995; Rubin, 1976). These reasons are often referred to as missing data mechanisms. Missing data mechanism is the process by which the data become missing.

There are three main classes of missing data mechanism. These are missing completely at random (MCAR), missing at random (MAR), and not missing at random (NMAR)(see Molenberghs and Fitzmaurice in Fitzmaurice et al., 2008, Ch. 17, pp. 307-400). These classifications/definitions are due to Rubin (Rubin, 1976, 1996). The MCAR mechanism assumes that the probability that a response will be missing is independent of the observed and the unobserved values. The MAR mechanism assumes that the probability that a response will be missing is independent of the unobserved values whereas NMAR mechanism assumes that the probability that a response will be missing depends on the observed and unobserved values.

In principle, when analyzing incomplete longitudinal data, a decision has to be made concerning the modeling approach for the missing data process. In the statistical literature for methods of handling missing data (Molenberghs et al., 2004; Molenberghs and Kenward, 2007; Verbeke and Molenberghs, 2012), the last observation carried forward (LOCF) and baseline observation carried forward (BOCF) have been severely criticized for their lack of credibility and statistical power, restrictive unrealistic assumptions, and inability to address relevant scientific questions (Molenberghs and Kenward, 2007; Carpenter et al., 2013) such as estimating the effect of treatment assuming that dropouts in the active treatment arm would exhibit similar statistical behaviour to subjects placebo arm. Also, these analyses cannot address scientific questions because the imputed or estimated values for a subject's missing responses are identical and hence does not account for uncertainty inherent in the imputation of the missing values (Little, 1995; Rubin, 1996).

The LOCF approach replaces missing values for dropouts in the study with their respective last observed values. This approach assumes that the response of a patient does not change after he or she drops out after randomization. Advocates of the LOCF method argue that it produces conservative estimates of treatment effect. Critics of LOCF argue against its conservatism, noting that the LOCF can exaggerate the magnitude of treatment effects and inflate Type I error, that is, falsely conclude a difference exists when in fact the difference is zero (Molenberghs et al., 2004; Molenberghs and Kenward, 2007). Molenberghs et al. (2004) showed that the LOCF is neither valid under general assumptions nor based on statistical principles. These authors concluded that it is not a sensible method, and should not be used (Carpenter et al., 2007).

The BOCF approach, on the other hand, replaces the missing values for each dropout in the study with their respective baseline observations. This approach assumes that responses for dropouts return to the same level as they were measured before randomization. This assumption is very intuitive in a clinical studies such as those involving HIV-positive patients or patients with chronic cancer (Carpenter et al., 2013). However, this assumption might not always hold, since most patients in a given study might have obtained significant benefit from the intervention after randomization and would not necessarily lose all the benefit while on treatment. The problem of missing data has been an active area of research in biostatistics, and there have been many advances in statistical theory and in our ability to implement that theory. Advances in statistical theory and in computer hardware and software have made many alternative methods straightforward to implement. Alternative imputation approaches to the LOCF and BOCF are the multiple imputation methods (Rubin, 1976, 1996; Carpenter et al., 2006, 2013; Ayele et al., 2014). The multiple imputation (MI) approach creates a set of K imputed values for the missing values thereby creating K-complete data sets. Unlike the LOCF and BOCF methods, MI accounts for the uncertainty inherent in the imputation of the missing data. Thus, the MI has the advantage of handling nonresponse in large data sets because it moves missing data burden from data analysis to data procedure. This approach has been shown to be attractive because it can be highly efficient, even for small values of K (Rubin, 1976, 1977, 1996; Carpenter et al., 2013).

An alternative approach, to missing values under MAR or MCAR, which proved to be efficient, is the likelihood -based methods (Rubin, 1976; Molenberghs and Kenward, 2007). Likelihood-based methods can produce valid statistical inferences in a certain sense that maximum likelihood estimation implicitly allows the missing values to be validly predicted using the observed data and a correct model for the joint distribution of the observed measurements (Fitzmaurice et al., 2008, Ch. 17, pp. 399).

In analyzing incomplete longitudinal data, valid likelihood inference may be obtained by assuming MAR or MCAR mechanism for the missing data. However, one cannot ignore the possibility of the data being not missing at random (Molenberghs et al., 1997; Molenberghs, Verbeke, Thijs, Lesaffre and Kenward, 2001; Molenberghs et al., 2008). When data are NMAR, the missing data mechanism has to be modeled implicitly or explicitly to obtain valid inferences. Molenberghs et al. (2008) pointed out that different models for data with an NMAR structure can lead to identical fits of the data but with different implications for the unobserved data. This contradiction may lead to different conclusions. Sensitivity analysis is therefore required to deal with this ambiguity in the modeling of data not missing at random.

The NMAR methods for sensitivity analysis has a long history (Rubin, 1976; Little, 1993; Diggle and Kenward, 1994; Verbeke et al., 2001; Diggle et al., 2002; Shen et al., 2006; Williamson, 2006; Molenberghs and Kenward, 2007; Molenberghs et al., 2008; Creemers et al., 2011, 2010). Sensitiv-

ity analysis is defined as an approach in which several statistical models are considered to explore whether conclusions are sensitive/robust to different assumptions about the missing data mechanism. It can also be defined as an approach where the statistical model is further examined using specialized tools such as diagnostic measures (see Molenberghs and colleagues in Fitzmaurice et al., 2008, Ch. 22, pp. 502-503). This is a more general definition which comprises a wide variety of methods. For instance, one way of conducting sensitivity analysis is to fit a selected number of NMAR models that make assumptions which deviates from the MAR assumption. Then, the level of robustness of statistical inferences under MAR provides and indication about the level of confidence that can be placed on the results. Smuk (2015) defined sensitivity analysis as an act of drawing conclusions under working assumptions about the missing data mechanism, identifying a set of plausible alternative assumptions, and examining the variation in statistical output and conclusions under the alternative setting.

The sensitivity analysis methods rely on NMAR joint models for non-ignorable missingness (Thijs et al., 2000; Verbeke et al., 2001; Shen et al., 2006; Molenberghs et al., 2008; Beunckens et al., 2008; Fitzmaurice et al., 2008; Molenberghs et al., 2014; Iddrisu and Gumedze, 2018). The NMAR is described as non-ignorable missingness since the missing data mechanism cannot be ignored, especially when the focus is to draw statistical inferences about the distribution of the complete longitudinal profile (see Molenberghs and Fitzmaurice in Fitzmaurice et al., 2008, Ch. 17, pp. 398-399). These models jointly model the measurement and the missingness processes. Some of the NMAR sensitivity analysis models used in this thesis have two-part specifications, where we specified a model for the measurement process and the model for the missingness process.

In this thesis, the NMAR models for sensitivity analysis assume the linear mixed effect model (Laird and Ware, 1982) for the measurement process. Whereas some of the NMAR models assume a logistic regression model for the probability of dropout (Diggle et al., 2002; Verbeke et al., 2001; Creemers et al., 2011), others (Carpenter et al., 2013; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Ayele et al., 2014) assumed an imputation model to obtain missing post-deviation outcome data. The NMAR sensitivity analysis methods (Wu and Carroll, 1988; Diggle and Kenward, 1994; Verbeke et al., 2001; Shen et al., 2006; Creemers et al., 2011) that specify a logistic regression model for the probability of dropout rely on choosing parameter values describing how post-deviation behavior evolves over time. This model is specified in such way that the probability of dropout at a given visit depends on only the previously observed and the current outcomes. The dropout model is specified in such a way that the response variable in the measurement model is a covariate in the dropout model. This specification of the dropout model has implication for sensitivity analysis. In particular, making different assumptions about the dependence of the dropout probability on the current response allows us to assess sensitivity of inference under such assumptions.

Wu and Carroll (1988); Thijs et al. (2000); Verbeke et al. (2001); Thijs et al. (2002); Tsonaka et al. (2009); Creemers et al. (2010) specified a logistic regression model for the probability of dropout where they assumed that the probability of dropout at a given time point depends on unobserved random effects. The measurement and the dropout model share common random effects. The shared random effects account for the association between the measurement and the dropout models as well the correlation between the repeated measurements. This specification has implications for sensitivity analysis. For instance, one can assume that the measurement and the dropout models do not share a common random effect on the one hand, and assume that the measurement and the dropout models shared common random effects on the other hand, and the compare the results.

Instead of specifying parameter values describing how post-deviation behaviour evolves, other methods Carpenter et al. (2013); Ratitch et al. (2013); Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf (2013); Mallinckrodt, Lin and Molenberghs (2013) identify groups of patients within the trial with this kind of behaviour, and then use these to construct the imputation distribution (Carpenter et al., 2013). This means that for each patient who deviates, the joint distribution of their pre- and post-deviation behavior is built. The joint distribution uses information from other groups of patients in the trial (the reference group) and then this information is used to construct the deviated patient's conditional distribution of post-deviation, given pre-deviation data. This conditional distribution of post-deviation data, given the pre-deviation data is in turn used to impute the patient's post-deviation data (Carpenter et al., 2013; Ratitch et al., 2013; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Permutt, 2015, 2016).

The NMAR models are often classified according to either the selection model (SeM), pattern-mixture model (PMM), or shared-parameter model (SPM) frameworks. In the selection model (SeM) framework, the joint distribution of the measurement and the dropout processes is factored as the marginal distribution of the measurement process and the conditional distribution of the dropout process, given the measurements (Rubin, 1976; Diggle and Kenward, 1994; Little, 1995). The pattern-mixture model (PMM) is a reverse factorization of the SeM defined as the marginal distribution of the dropout process and the conditional distribution of the measurement process given the dropout process (Rubin, 1976; Little and Yau, 1996).

For the shared-parameter model (SPM), a set of latent variables (random effects) is assumed to be shared between the measurement and the dropout processes (Rubin, 1976; Wu and Carroll, 1988; Tsonaka et al., 2009). It is conventionally assumed that, given this set of random effects, no further dependency exist between the measurement and the dropout process, although this can be generalized (Creemers et al., 2011, 2010). Yuan and Little (2009) proposed mixed-effect hybrid models (MEHMs) framework, where the joint distribution of the measurement process and dropout process is factorized into the marginal distribution of random effects, the dropout process

conditional on random effects, and the measurement process conditional on dropout patterns and random effects.

The definition of MAR is in SeM terms (Little, 1995; Rubin, 1976; Creemers et al., 2011, 2010). Molenberghs et al. (2008) provided a characterization for the PMM, whereas Creemers et al. (2010, 2011) characterized the MAR in the SPM framework using the extended SPM framework. They showed how the SPM can be constrained such that the dropout at a given point in time can depend on current and past measurements but not future measurements.

The NMAR models appear to offer the most suitable framework for handling missing data. However, care is required when one is interpreting evidence for or against NMAR using only the data under study. This is because for every NMAR model, there is an MAR counterpart with exactly the same fit to the observed data but with differing interpretation for the observed data (Molenberghs et al., 2008). Molenberghs et al. (2008) noted that a compromise between placing too much confidence in such an NMAR model on one hand or ignoring them completely on the other hand, consists of making them a component of sensitivity analysis. The sensitivity analysis is the main focus of this thesis. Our sensitivity analysis are based on the selection model (Diggle and Kenward, 1994; Verbeke et al., 2001), the pattern-mixture model (Carpenter et al., 2013; Mallinckrodt, Lin and Molenberghs, 2013), and the shared-parameter model (Tsonaka et al., 2009; Creemers et al., 2011, 2010) frameworks.

1.1 Contributions

In this thesis, we have conducted sensitivity analyses for incomplete longitudinal data and have made the following contributions to knowledge and research:

1. We have implemented sensitivity analyses approaches that can be contextually applied to the IMPI trial and other trials with similar settings. This is important because the objective of a trial varies from trial to trial and the method of analysis depends on the trial's objective.

2. We have performed sensitivity analysis to investigate for dropout probability which is likely to overturn treatment effect. Such analyses are vital for medical researchers and clinicians to stress-test (over-stretched) missing data assumptions to assess their impact on the statistical inferences and to offer clinical explanations associated with such results. The results from these approaches as well as a SAS code for their implementation have been published Iddrisu and Gumedze (2018) paper (see Chapter 4).

3. The pattern-mixture with model imputation (PM-MI) approach was proposed by Carpenter et al. (2013) and applied to measurements at the last visit (assuming that measurements are

independent). So they use the linear regression model that assumes that the measurements are independent. In this thesis, we considered the PM-MI for longitudinal measurements, (measurement at all visits), assuming that such measurements are correlated using the linear mixed effects model for the correlated measurements. We also conducted simulation studies to evaluate the performance of the sensitivity analysis approaches (Iddrisu and Gumedze, 2019*a*).

4. We proposed and conducted sensitivity analysis for the shared-parameter model framework via the global influence approach. We have also conducted simulation studies to evaluate the performance of the shared-parameter models considered in this thesis (Iddrisu and Gumedze, 2019*b*).

1.2 Rationale

It is known that missing data may severely compromise statistical inferences from clinical trials. Hence, there is the need for regulatory guidelines on design, conduct, and analysis of clinical trials with missing data. The national research council (NRC, 2010) report outlined recommendations for the prevention and treatment of missing data in clinical trials. When drawing statistical inferences from incomplete data, methods of analysis based on plausible scientific assumptions should be used. However, these assumptions cannot be verified using the data at hand, and hence there is the need to conduct sensitivity analysis in order to investigate the robustness of statistical inferences to plausible alternative assumptions about the missing data. Sensitivity analysis to missing data is the main aim of the thesis and we will focus on development of sensitivity analysis strategies and their application to incomplete longitudinal CD4 count data from the IMPI trial (Mayosi et al., 2014, 2012). We will also design simulation experiment to investigate sensitivity of the sensitivity analysis strategies to varying rate of missing values and mis-specification. We also seek to provide recommendations on which sensitivity analysis strategy to use and in what circumstance or clinical trial setting.

1.3 Objectives

The objectives of this study are to

(i) develop sensitivity analysis strategies for incomplete longitudinal data and to apply these strategies to incomplete longitudinal CD4 count data from the IMPI trial.

(ii) design simulation experiment to investigate sensitivity of the sensitivity analysis strategies to different rates of missing data and mis-specification.

(iii) make recommendation on which sensitivity analysis approach can be used and in what clinical trial setting.

1.4 Outline of the thesis

In this chapter, we have given the introduction, rationale, and objectives of the thesis. This section also provides an outline of the thesis. This thesis is divided into seven chapters.

Chapter 2 provides discussion on longitudinal study design, missing data patterns, missing data mechanisms, standard methods for handling missing data, and standard method of analysis for longitudinal data.

In Chapter 3, we review the literature on joint models for non-ignorable missing data (NMAR) (see Molenberghs and Fitzmaurice in Fitzmaurice et al., 2008, Ch. 17, pp. 401). Some of the main joint modeling frameworks reviewed are the selection model (Diggle and Kenward, 1994), pattern-mixture model (Little, 1993), and shared-parameter model (Follmann and Wu, 1995*a*) frameworks. This review is necessary for a proper understanding of the missing data assumptions associated with such specifications.

In Chapter 4, we discuss the Diggle and Kenward's (DK) selection model (Diggle and Kenward, 1994) to nonrandom missing data. This discussion is followed by a discussion on the application of local influence approach to the DK selection model (Kenward, 1998; Thijs et al., 2000; Verbeke et al., 2001; Shen et al., 2006). We then introduce the stress-testing methodology which allows us to assess the effects of extreme dropout probability assumptions on treatment effect and dropout mechanism. These approaches were implemented to the CD4 count data and results were discussed.

In Chapter 5, our sensitivity analyses (Carpenter et al., 2013) are based on the pattern-mixture model framework (Wu and Bailey, 1988, 1989; Wu and Carroll, 1988). We introduce the pattern-mixture model and then focus on the pattern-mixture with multiple imputation approach (Carpenter et al., 2013). We apply theses approaches to the CD4 count data from the IMPI trial and then discuss the results. We also performed simulation experiment to evaluate the performance of the PM-MI approach.

In Chapter 6, the sensitivity analyses are based on the conventional shared-parameter model framework (SPM) (Wu and Bailey, 1988; Wu and Carroll, 1988; Tsonaka et al., 2009). We introduces the SPM framework and then focus on the generalized shared-parameter model (GSPM) (Creemers

et al., 2010, 2011). We also propose to assess the effect of potentially influential subjects on parameters estimates, under the shared-parameter model, using the global influence approach. We applied these approaches to the CD4 count data from the IMPI trial and discussed results.

In Chapter 7, we discuss and draw conclusions, and also identify areas of further research.

Chapter 2

Fundamental concepts of incomplete longitudinal data

2.1 Introduction

In this chapter, we discuss the general structure of longitudinal data and incomplete longitudinal data. This chapter also highlights some of the issues associated with incomplete longitudinal data. We then discuss methods for handling missing data as well as standard methods of analysis for such data. This is necessary to gain insight into some of the common problems associated with incomplete longitudinal data and suitable methods for analyzing such data.

We commence with a detailed review of the general structure of the longitudinal data, missing data patterns and the missing data mechanisms (Rubin, 1976, 1987; Little and Yau, 1996). We then discuss the three key recommendations by the NRC panel (NRC, 2010) on prevention and treatment of missing data in clinical trials. We also review the methods for handling missing data as well as how to draw inference from them. For the methods for handling missing data, we review the simple imputation (SI) (Molenberghs and Kenward, 2007), multiple imputation (MI) (Rubin, 1996), and the direct maximum likelihood (ML) (Dempster et al., 1977) methods. We compare the MI and the ML methods. This comparison highlights some advantages and disadvantages associated with using theses methods. This thesis focuses on likelihood-based methods of analysis and methods based the generalized estimating equations (GEE) (Liang and Zeger, 1986) and the weighted GEE (wGEE) (Robins et al., 1995; Robins and Rotnitzky, 1995; Scharfstein et al., 1999; Seaman and White, 2013) are not considered.

For the standard methods of analysis for longitudinal data, we present the general form of the linear mixed model (LMM) (Laird and Ware, 1982). Detailed information on how to estimate parameters, in the LMM, can be found in Laird and Ware (1982); Henderson et al. (1959); Patterson and Thompson (1971); Harville (1974); Searle (1995); Gumedze (2008); Searle et al. (2009); Verbeke and Molenberghs (2009). Thereafter, we introduce the IMPI trial data with focus on the CD4 count measurements (Mayosi et al., 2012, 2014).

2.2 Longitudinal data structure

Longitudinal data refers to study designs that include measurements for the same response variable taken on several occasions for each patient. Such designs result in a "response profile" for each patient. The aim is to model and compare the mean response profiles for different groups or strata. The groups are defined by a main exposure variable, such as treatment. The dimension over which the repeated measurements are taken is often time.

We assume that for each of N patients, we intend to record n repeated measures of the response variable on the same patient. It follows that a subject with a complete set of responses has an $n \times 1$ response vector denoted by $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{in})'$. The response Y_{ij} is measured for each subject $i, i = 1, \ldots, N$ at occasion $j, j = 1, \ldots, n$. Also, associated with $\mathbf{Y}_i$ is an $n \times p$ matrix of covariates, $\mathbf{X}_i$. We have explained in Chapter 1 that it is common to have some components of $\mathbf{Y}_i$ missing in a longitudinal clinical study designs. Let $\mathbf{R}_i$ be an $n \times 1$ vector of response indicators, $\mathbf{R}_i = (R_{i1}, \ldots, R_{in})'$, with $R_{ij} = 1$ if Y_{ij} is observed and $R_{ij} = 0$ Y_{ij} is missing. Given $\mathbf{R}_i$, the complete data $\mathbf{Y}_i$ can be partitioned into observed $\mathbf{Y}_i^o$ and missing $\mathbf{Y}_i^m$ components. The response indicator $\mathbf{R}_i$ is recorded for all patients. The likelihood function is defined as $\Pr(\mathbf{Y}_i, \mathbf{R}_i)$ and the observed data likelihood function is $\Pr(\mathbf{Y}_i^o, \mathbf{R}_i)$, where Pr denotes probability.

In model building, the interest is to specify a density function for the data $\Pr(\mathbf{Y}_i, \mathbf{R}_i)$ and then investigate how $\mathbf{Y}_i$ depends on $\mathbf{X}_i$ over time. This require us to incorporate missing data into the statistical analysis.

2.2.1 Missing data patterns

In this section we discuss some of the missing data patterns that are common in longitudinal clinical trials. There are several classes of missing data patterns. Some of the common classes of overall missing data patterns include (1) univariate, (2) monotone, (3) file matching, and (4) general pattern of missingness (Molenberghs et al., 2014, pp.270). Figure 2.1 illustrates the missingness patterns using four artificial datasets (Van Buuren, 2007, 2012). The x-axis represents patient i and the y-axis represents the schedule visit, whereas the blue and brown colours refer to observed and missing responses respectively. Each square refers to group of patients which are observed (indicated by blue square) and missing (indicated by the brown square).

For univariate pattern, missing values occur on one patient or group of patients that are entirely observed or missing, whereas in monotone pattern, patients are ordered such that if patients i is missing at scheduled visit time j, then that same patients i will be missing at the scheduled visit time $j+1, \ldots, n$, where n denotes the last scheduled visit. A typical scenario of file matching pattern

of missingness occurs when data sets from different studies are merged together for analyses. This leads to some common variables being fully observed in all data sets while some variables are only observed in individual data sets and therefore missing in others. General or intermittent pattern of missingness on the other hand occurs when missing values are randomly scattered.

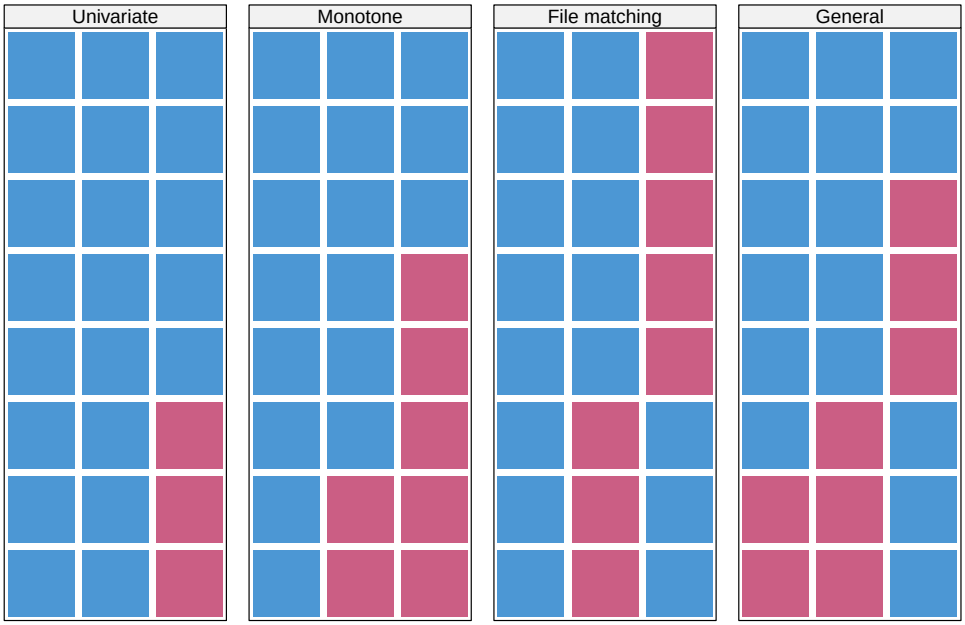

Figure 2.1: Illustration of missing data patterns

2.2.2 Missing data mechanisms

This section reviews the general taxonomy of missing data mechanism, originally introduced by Rubin (1976). These missing-data mechanisms differ in terms of assumptions about whether missingness is related to observed and unobserved response data. This section also discusses the implications of these missing-data mechanisms for analysis of longitudinal data. These assumptions and their implications are vital for both primary analysis and for framing of relevant sensitivity analyses (Verbeke et al., 2001; Creemers et al., 2011; NRC, 2010; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Carpenter et al., 2013).

The original work of Rubin (1976) provided a formal framework for studying incomplete longitudinal data by categorizing reasons for missingness into three major taxonomies. These taxonomies include (1) missing completely at random (MCAR) (2) missing at random (MAR) and (3) not missing at random (NMAR).

2.2.2.1 Missing completely at random (MCAR)

Incomplete longitudinal data are described as MCAR when the probability that responses $\mathbf{Y}_i$ are missing is independent of the observed $\mathbf{Y}_i^o$ and unobserved $\mathbf{Y}_i^m$ components of $\mathbf{Y}_i$. Suppression covariates and parameters, this type of missing data mechanism can alternatively be defined as $\Pr(\mathbf{Y}_i \mid \mathbf{R}_i) = \Pr(\mathbf{R}_i)$. For instance, MCAR may occur in a longitudinal study design that allows individuals to move in and out of the study after providing a predetermined number of repeated measures. This is often done so as to reduce response burden (Fitzmaurice et al., 2008, Ch. 17, pp.396-400). With such a design, it can be observed that measurements depend on the study design and not in any way related to the response variable. This is because, the timing and the number of the measurements are based on the study design. In addition, Ratitch et al. (2013) noted that data missing due to administrative reasons could be classified as MCAR since the reason for missingness has nothing to do with the response model or the covariates.

The observed data $\mathbf{Y}_i^o$ under MCAR assumption can be regarded as a random sample from complete data. Hence, all inferences from the observed data coincide with the corresponding inferences from the joint distribution of the complete data $\mathbf{Y}_i$ (Fitzmaurice et al., 2008, pp.397). The implications of such an assumption are that (1) the completers can be regarded as a random sample from the target population, (2) the conditional distribution of the observed data $\mathbf{Y}_i^o$ for those subjects with any non-response pattern coincide with the same component of the response variable $\mathbf{Y}_i$ in the target population, and (3) the distribution of $\mathbf{Y}_i^m$ for those subjects with any non-response pattern coincides with the same component of the response variable $\mathbf{Y}_i$ for the completers. These imply that any method of analysis that yields valid inferences in the absence of missing data will also yields valid inferences when the analysis is confined to the completers. Such analysis is often referred to as complete-case (CC) analysis (Molenberghs and Kenward, 2007).

2.2.2.2 Missing at random (MAR)

Incomplete longitudinal data are described as MAR when the probability that responses are missing depends on the observed response $\mathbf{Y}_i^o$, but given the observed data, is independent of the unobserved $\mathbf{Y}_i^m$ components of the response variable $\mathbf{Y}_i$. Alternatively, we can define MAR as $\Pr((\mathbf{R}_i \mid \mathbf{Y}_i) = \Pr(\mathbf{R}_i \mid \mathbf{Y}_i^o)$. This type of missing data mechanism can be thought of as having a logistic form. That is, MAR may occur in a longitudinal study where the study protocol stipulates that a patient should be removed from the study if such patient's responses does not meet certain specified requirements set in the study protocol (Fitzmaurice and Laird, 2000, pp. 399-400). In such a longitudinal study design, missingness in the $\mathbf{Y}_i$ is under the control of the researcher and hence, the probability of missingness depends on only $\mathbf{Y}_i^o$.

One implication of the MAR assumption is the "completers" are a biased sample from the target population and hence, an analysis restricted to the "completers" is not valid. It follows that (1) the observed data can no longer be regarded as a random sample from the complete data, (2) completers can no longer be regarded as a random sample from the target population, and (3) conditional distribution of the $\mathbf{Y}_i^o$ for patients with any non-response pattern does not coincide with the distribution of the same component of $\mathbf{Y}_i$ in the target population.

However, MAR assumption provides us with an important information about the distribution of the missing data. That is, the conditional distribution of a patient's missing data $\mathbf{Y}_i^m$, given the observed data $\mathbf{Y}_i^o$, is the same as the conditional distribution of the corresponding observations for the "completers", conditional on the "completers" having the same values as $\mathbf{Y}_i^o$. As opposed to MCAR, under MAR, the validity of assumptions about the missing data cannot be justified from the data at hand without assuming an alternative model for the missing data.

2.2.2.3 Not missing at random (NMAR)

Incomplete longitudinal data are described as NMAR mechanism when the probability that responses are missing depends on the observed $\mathbf{Y}_i^o$ and the unobserved $\mathbf{Y}_i^m$ components of the response variable $\mathbf{Y}_i$. Alternatively, NMAR can be defined as $\Pr(\mathbf{R}_i \mid \mathbf{Y}_i) = \Pr(\mathbf{R}_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m)$. Since missingness cannot be ignored, this type of missing data mechanism is often referred to as non-ignorable missingness (Molenberghs et al., 2008; Fitzmaurice et al., 2008, Ch.17, pp.401-403). This means that any valid inferential method requires specification of a model for the missing data mechanism. The implications are that the distribution of $\mathbf{Y}_i^m$, conditional on $\mathbf{Y}_i^o$ and any potential covariates $\mathbf{X}_i$, no longer coincides with the distribution of the completers in the target population. Rather, the distribution of $\mathbf{Y}_i^m$ depends upon $\mathbf{Y}_i$ and on $\Pr(\mathbf{R}_i \mid \mathbf{Y}_i, \mathbf{X}_i)$. Therefore, a model assumed for $\Pr(\mathbf{R}_i \mid \mathbf{Y}_i, \mathbf{X}_i)$ is crucial, must be correctly specified, and included in the analysis in order to produce valid inferences.

Under NMAR mechanism, the observed data provide no information that can be used to either substantiate or disprove one NMAR mechanism over another. Recognizing that without additional information, testing for such assumptions depends solely on such untestable assumptions, several authors (Little and Rubin, 1987; Laird, 1988; Molenberghs, Kenward and Goetghebeur, 2001; Verbeke et al., 2001; NRC, 2010; CHMP, 2010; Thabane et al., 2013) placed emphasis on non-ignorable missing data methods for conducting sensitivity analysis.

At this point, it is important to note that MCAR, MAR, and MNAR are assumptions made regarding the underlying non-response process. As such, absolute certainty about them can never be established. Incompleteness in the data induces a certain degree of unidentifiability. This unidentifiability has to be compensated for by a careful choice for the operating missingness mechanism.

Though an experienced researcher may have some general idea as to the nature of the missing data, this would be, at best, an educated hypothesis. Challenged with unfamiliarity with clinical (and other) aspects of non-response, one can perhaps avoid the more implausible MCAR, or even go for some sort of sensitivity analysis, comparing results under different mechanisms (Molenberghs et al., 2008).

The validity of inferences made under different statistical methods depends primarily on the assumed missingness mechanism. In extreme cases, contradictory results can be obtained under two different mechanisms (Sotto, 2009). Moreover, a model obtained under an MNAR missingness mechanism admits an equivalent MAR model, at least in terms of the models which fit to the observed data, but with contrasting inferences regarding the unobserved data (Molenberghs et al., 2008; Sotto, 2009). Hence, sensitivity analysis is necessary to to ascertain the nature of the missing data mechanism operating under a given clinical trial. Such ideas will be discussed in greater detail in sensitivity analyses, Chapters 4-6.

2.3 Common missing data issues and solutions

Following the report released by the National Research Council (NRC) of the National Academy of Sciences entitled "The Prevention and Treatment of Missing Data in Clinical Trials", commissioned by the US Food and Drug Administration (FDA), the report (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; NRC, 2010) had an immediate effect on the way in which statisticians and clinical researchers in both industry and regulatory agencies think about the missing data problem (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Carpenter et al., 2013; Ratitch et al., 2013; Permutt, 2015, 2016). This report outlined recommendations which form the basis for designing clinical trials as well as conducting analysis that has great potential to improve study quality and the way in which results of the analysis can be interpreted. This can be achieved by reducing the amount of missing data through changes in trial design and conduct, and by planning and conducting analyses that better account for the missing information. When data are missing, validity of any methods of analysis will depend on the scientific question addressed in a given clinical trial setting. This is because each trial may have different settings as well as varying scientific questions to address. In this thesis, we describe some of the recommendations in the report and discuss how these recommendations are addressed using the CD4 count data from the IMPI trial (Mayosi et al., 2012, 2014).

Randomized clinical trials are the recommended tool for evaluating the effect of new medical interventions. Randomization provides for a balance comparison between treatment and control groups, balancing out, on average, distributions of known and unknown factors among trial participants

(NRC, 2010). However, a substantial percentage of the measurements of the outcome or outcomes of interest is often missing. This missingness reduces the benefit provided by the randomization and introduces potential biases in the comparison of the treatment groups (NRC, 2010).

The presence of missing data can arise due to a variety of reasons, including the inability or unwillingness of participants to meet appointments for evaluation. Some of the reasons could be due to the adverse effect of treatment, or moving to partial compliance with treatment (NRC, 2010; Carpenter et al., 2013). The NRC panel noted that the existing guidelines for the design and conduct of clinical trials, and the analysis of such data, provide only limited advice on how to handle missing data. Hence, approaches to the analysis of data with a substantial number of missing values tend to be ad hoc and variable. Consequently, the panel concludes that a more principled approach to design and analysis in the presence of missing data is both needed and possible. The panel noted that this approach needs to focus on two critical elements. The first is to carefully design and conduct trials to limit the amount and impact of missing data. This requires a trial design to clearly define the target population, and the outcomes that will form the basis for decisions about efficacy and safety. The treatment of missing data depends on how these outcomes are defined, and lack of clarity in their definition translates into a lack of clarity as to how to deal with missing data issues.

Given the difficulties of adequately addressing missing data at the analysis stage, the design process needs to pay more attention to the potential hazards arising from substantial numbers of missing values. The recommendation 2 of the NRC panel states that "investigators, sponsors, and regulators should design clinical trials consistent with the goal of maximizing the number of participants who are maintained on the protocol-specified intervention until the outcome data are collected ". The second approach is to conduct analysis that makes full use of information on all randomized participants and is based on careful attention to the assumptions about the nature of the missing data underlying estimates of treatment effects.

Among the recommendations by the NRC panel are three key recommendations. These recommendations are (1) making precise and clear objectives of the trial, (2) minimizing the amount of missing data, and (3) using plausible primary analysis together with sensitivity analyses that support the research hypotheses to be addressed, as well as being capable of assessing sensitivity of primary results to missing data assumptions. The NRC (2010) recommendations have received much attention. In particular, clinical practitioners, academicians and regulators now require drug development groups to be guided by such recommendations when proposing and implementing plans to deal with missing data. Although the NRC panel's work focused primarily on Phase III confirmatory clinical trials that are the basis for the approval of drugs and devices, many of the panel's recommendations can be applied to all randomized trials (NRC, 2010, pp. 1).

We discuss the three key recommendations of the NRC panel and then point out how these recommendations are applied in the IMPI trial settings and trials with similar settings.

2.3.1 Establish clear and precise objectives of the trial

The NRC (2010) panel recommends to set out clear and precise objectives of the trial. This is a measure to avoid ambiguities in conclusions caused by missing data. The data may be missing intermittently or missing because of dropout. Depending on the trial's objectives, dropouts may or may not be given standard of care after dropout. Assessment of dropout from the initially randomized treatment or introducing a standard of care, may or may not be considered in the primary analysis. The primary analysis (as specified in the statistical analysis plan) addresses the main objective of the study. It is important to be aware that dropout analysis only occurs when patients deviate from the initially randomized treatment (and either discontinue treatment or switch to standard of care) and observations are made but not used in the analysis.

The use of post-deviation data largely depends on the estimand. The post-deviation data are data obtained for subjects after dropout. Within a single study, some analyses may use follow-up data while others may not. The debate on appropriate estimands is based on whether the focus is on efficacy or effectiveness (Carpenter et al., 2013; Mallinckrodt, Lin and Molenberghs, 2013; Ayele et al., 2014). For instance, if one is interested in the difference in outcome improvement at the planned end period for all randomized patients, post-deviation data can be used in the primary analysis. In this scenario, the hypothesis to address is the effectiveness estimand hypothesis. This is because the effectiveness estimand compares treatment groups irrespective of what treatment patients received, and thus inferences is on the effectiveness of the treatment regimen and not the originally randomized treatment. On the other hand, if one is interested in the difference in outcome improvement assuming that all patients adhered to treatment, then post-deviation data cannot be used in the primary analysis. In this scenario, the efficacy estimand will be of interest since the estimand compares causal effects of the initially randomized treatment if taken as directed.

Furthermore, if one is interested in the difference in outcome improvement in all randomized patients at the planned end period of the trial attributable to the initially randomized treatment, post-deviation or imputed data can be used. In this case the hypothesis to address is the effectiveness hypothesis. Whenever there is the need to use post-deviation data, control imputation analyses may be used as a means to obtained follow-up data needed to estimate effectiveness (Carpenter et al., 2013; Mallinckrodt, Lin and Molenberghs, 2013; Ayele et al., 2014).

It can be observed under the data description Section 2.6 that our hypothesis to address is the effectiveness hypothesis. This is because the trial aim is on outcome improvement at the planned end period for all randomized patients.

2.3.2 Minimizing the amount of missing data

The best approach to missing data is to avoid the occurrence of missing data. However, missing data are often unavoidable since the missingness are often not under the control of the researcher, especially studies involving human subjects. The development of new analytic methods and software tools for analyzing incomplete data has been an active area of research and more achievements have been made in that regards. However, all analyses are still challenged by the confusing and difficult problem of analyzing incomplete data (NRC, 2010). For instance, all analyses require assumptions about the missing data; these assumptions cannot be verified from the data, and the appropriateness of analyses and inference cannot be ensured.

When considering design options that minimize missing data, the influence that these options may have on other aspects of the trial must be considered. According to NRC (2010), some of the trial design options that one could consider include enrolling a target subpopulation for whom the risk-benefit of the treatment is more favorable, or identifying such subgroups during the course of the trial via enrichment or run-in designs. However, examples where this has been done successfully in the context of lowering rates of dropout are rare. Other design options in the NRC guidelines included use of add-on designs (Chow and Liu, 2008) and flexible dosing (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013). One may also minimizes patient burden by using efficient data capture procedures, providing education on the importance of complete data, monitoring and providing incentives for complete data.

The missing data methods discussed under the Section 2.4 address this need by accounting for the missing data in the statistical analysis. The validity of these methods lies on the objective of the trial as stated under the recommendation 2.3.1. By the recommendation 2.3.1 and in the context of the IMPI clinical trial, the use of post-deviation data, which we do not have, can be compensated for by using methods that account for, or estimate the distribution of missing values before data analysis.

2.3.3 Appropriate primary and sensitivity analyses

Among the NRC key recommendations is the use of an appropriate primary analysis and sensitivity analyses approaches to assess the sensitivity of the primary analysis results to key assumptions about the missing data. Both primary and sensitivity analysis are necessary because despite all efforts to minimize missing data, anticipating complete data is not realistic. In order to decide on appropriate primary analysis, the process generating the missing data must be take into account in the statistical analysis. In other words, one must consider the missing data mechanisms (MCAR, MAR and NMAR) in order to decide on appropriate primary analysis to use (Carpenter et al.,

2013).

The statistical analysis methods such as maximum likelihood (Dempster et al., 1977; Harville, 1977), multiple imputation (Lavori et al., 1995; Rubin, 1996; Carpenter and Goldstein, 2004; Ayele et al., 2014), Bayesian methods (Harville, 1974; Daniels and Hogan, 2008), and methods based on weighted generalized estimating equations (Robins et al., 1995; Robins and Rotnitzky, 1995; Scharfstein et al., 1999; Seaman and White, 2013) can reduce the potential bias arising from missing data by making principled use of auxiliary information available for nonrespondents. These methods assume that the data are MAR. The NRC panel encourages increased use of these methods. However, these methods rely on untestable assumptions concerning the factors leading to the missing values and how they relate to the study outcomes. Therefore, the assumptions underlying these methods need to be clearly communicated to medical experts so that they can assess their validity (Carpenter et al., 2013).

Sensitivity analyses are therefore important to assess the degree to which the treatment effects rely on the assumptions considered. We need to choose the primary analysis approach carefully since it is based on the chosen primary analysis method that appropriate sensitivity analysis can be formulated to assess sensitivity of the results under the primary analysis to the sensitivity analysis assumptions (Diggle and Kenward, 1994; Thijs et al., 2000; Verbeke et al., 2001; Thijs et al., 2002; Shen et al., 2006; Creemers et al., 2011; Mallinckrodt, Lin and Molenberghs, 2013; Carpenter et al., 2013).

In this thesis, the primary analysis model assumes that the data are missing at random and sensitivity analyses are based on the selection model (Diggle and Kenward, 1994; Verbeke et al., 2001; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013), pattern-mixture model (Wu and Bailey, 1989; Carpenter et al., 2013; Mallinckrodt, Lin and Molenberghs, 2013), and the shared-parameter model (Gao, 2004; Tsonaka et al., 2009; Creemers et al., 2011, 2010).

2.4 Drawing inferences from incomplete data

In this section, we briefly review some of the approaches for drawing inferences from incomplete data. We review some of the commonly-used methods for handling missing data in longitudinal study designs. For a number of commonly-used methods, users are not always aware of the assumptions that underlie the methods and the results drawn from applying them (Molenberghs and Kenward, 2007; Carpenter and Kenward, 2012; Carpenter et al., 2007). This lack of awareness is particularly true of single imputation methods such as last or baseline observation carried forward and random effects (mixed effects) regression models that rely on strong parametric assumptions.

When non-trivial proportion data are missing, a decision has to be made concerning the modelling approach. There is no universal method for handling incomplete data in a clinical trial. Each trial has its own set of design and measurement characteristics. However, the following are some sets of principles that can be applied in a wide variety of settings (NRC, 2010; CHMP, 2010; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013).

First, there is the need to determined whether missingness of a particular value hides a true underlying value that is meaningful for analysis. This may seem obvious but is not always the case. For instance, consider a longitudinal analysis of CD4 count data in a clinical trial for AIDS (NRC, 2010). For subjects who dropped out from the study because they moved to a different location, it makes sense to consider the CD4 counts that would have been recorded if they had remained in the study. For subjects who die during the course of the study, it is less clear whether it is reasonable to consider CD4 counts after time of death as missing values.

Second, it is important to document reasons for missing data as much as possible. This includes full and detailed documentation for each individual of the reasons for missing records or missing observations and particularly how this relates to the value of the outcome measure (NRC, 2010). Third, the trial-designers should decide on a primary set of assumptions about the missing data mechanism. Those primary assumptions then serve as a reference for the sensitivity analyses. In many cases, the primary assumptions can be missing at random (MAR). Assumptions about the missing data mechanism must be transparent and accessible to clinicians.

The fourth principle is that trial sponsors should conduct a statistically valid analysis under the primary missing data assumptions. If the assumptions hold, a statistically valid analysis yields consistent estimates, and standard errors and confidence intervals account for both sampling variability and for the added uncertainty due to missing observations. Fifth, it is important for the analysts to assess the robustness of the treatment effect inferences under the primary analysis by conducting a sensitivity analysis. The sensitivity analysis should relate treatment effect inferences to one or more parameters that capture departures from the primary missing data assumption (NRC, 2010; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013).

While other missing data methods may model missing data explicitly, other methods model missing data implicitly. There are three main methods for dealing with incomplete data. These are the (1) imputation methods, (2) likelihood-based methods, and (3) weighting methods. In general, these methods are interrelated in such a way that they "impute" certain values for missing data. The only difference is that the imputation methods impute values for missing values explicitly, whereas the likelihood-based and the weighting methods "impute" missing values implicitly. In the following sections, we point out the missingness mechanism required for each of these methods to yield valid statistical inferences. The advantages and disadvantages associated with using such methods are

discussed.

2.4.1 Complete-case analysis

First, we discuss the complete-case (CC) analysis. The CC analysis is not a method for handling missing data. This analysis confines analysis to only the completers. Completers are patients with no missing values. The CC analysis is valid when the mechanism of the missing data is MCAR. Statistical analyses can be performed using standard methods of analysis. This method is straightforward to use. Some of the disadvantages associated with using such analysis are (1) small sample size if there is substantial missing values, (2) loss of statistical power due to the reduction in the sample size and (3) biased inferences. Kim and Curry (1977) noted that the CC method may be preferred under a situation where the (1) sample size is large, (2) proportion of missing data is small, and (3) missing data mechanism is MCAR. Zhu (2014) noted that in the case of MCAR mechanism, the CC method will yield unbiased parameter estimates with larger standard errors because of the smaller sample size. This implies that MCAR assumption does not guarantee high precision (small standard errors) since there is potential loss of essential information due to missing data.

The CC analysis reduces efficiency such that the deviation around the true estimate is too large, and severe bias may occur if data are not MCAR (Little, 1992; Greenland and Finkle, 1995; Schafer and Graham, 2002). Severe bias can result when the missingness mechanism is MAR but not MCAR. This bias can be positive or negative, as illustrated by (Molenberghs et al., 2004). This approach may yield reliable estimates of treatment effect in a clinical trial setting where patients remain on treatment but investigators or clinicians could not measure their response to treatment at all scheduled visits time, due to administrative reasons (eg., inadequate resources).

Rather than deleting patients with missing values and then restricting analysis to only the completers, one can also consider analysis that used all the observed data. This can be achieved by using the linear mixed effect model (Laird and Ware, 1982). This model uses all the observed from both complete patients (patients with no missing data) and incomplete patients (patients with missing data).

2.4.2 Imputation-based approaches

Imputation of missing data is one of the approaches for handling missing data. The basic idea behind the imputation approach is that missing values are filled with the imputed values to obtain complete data sets. One advantage of the imputation approach is that, once a filled-in data set has been obtained, standard methods for complete data can be applied. There are two main and

widely-used imputation methods. These are the simple and multiple imputation methods. These methods model the missing data explicitly.

2.4.2.1 Simple imputation methods

Simple imputation (SI) methods explicitly impute a single value for the missing data and replace all unobserved responses with such value. Simple imputation approaches are referred to as single imputation approaches, since each missing value is only imputed once. Sometimes such methods are collectively termed as "Ad-hoc" because they are not derived from statistical principles. There are two widely used simple imputation methods. These methods are (1) the last observation carried forward (LOCF) and (2) the baseline observation carried forward (BOCF).

2.4.2.1.1 Last observation carried forward method

The last observation carried forward method has received a lot of attention (Molenberghs and Kenward, 2007). The LOCF approach replaces missing values of each patient in the study with their respective last observed values. This approach assumes that the response profile of a patient's does not change after he or she drops out after randomization. This approach is questionable in many settings since dropouts might have lost some treatment benefit they have obtained while on treatment (Molenberghs and Kenward, 2007). For this approach to make sense, very strong and often unrealistic assumptions have to be made. This is because at times, one has to believe that a patient's measurements stay at the same level from the time of dropout onwards or during the period they are unobserved. In other clinical trial settings, one might believe that the response profile changes as soon as a patient dropped out of treatment and even that it would flatten. This method (LOCF) shares with other single imputation methods the problem that it overestimates the precision by treating imputed and actually observed values on equal footing.

It is known (Molenberghs and Kenward, 2007, pp.41-42) that LOCF can be used for imputation of missing values, but need not be viewed as an imputation strategy, depending on the research question to address. Verbeke and Molenberghs (2012) showed that all components (fixed and random parts) of a linear mixed model (Laird and Ware, 1982) may be severely affected by using the LOCF method. A situation where interest focuses on the last-observed measurement, can often be considered to be a genuine motivation for LOCF. When this situation is considered plausible, the implication is that the problem of missingness is avoided and hence the LOCF method in such a situation cannot provide any evidence about the relative performance of methods used for handling missingness (Molenberghs and Kenward, 2007, pp.46).

Advocates of LOCF sometimes argue that it leads to conservative estimates of treatment effects

(Molenberghs and Kenward, 2007). However, it can be shown that this cannot be true in general (Molenberghs et al., 2004). Rather, the direction of the bias depends on the (unknown) true treatment effect and the missing data mechanism. In general, the LOCF is biased even when a complete case analysis is sensible (Molenberghs et al., 2004). In addition, the notion that LOCF produces conservative results does not appear to have arisen from formal proofs or rigorous empirical study. Molenberghs and Kenward (2007) showed that LOCF can exaggerate the magnitude of treatment effects and inflate the Type I error, that is, falsely conclude that a difference exists when in fact the difference is zero. Using data from the Isolde trial, Carpenter and Kenward (2007) showed that the LOCF is not sensible, even when data are MCAR. Carpenter et al. (2003) pointed out that the LOCF is unable to detect a wide range of actual treatment effects. They noted that in situations where the treatment effect can be detected by the LOCF procedure, the likelihood-based analyses will usually be more powerful.

The LOCF is neither valid under general assumptions, nor based on statistical principles. It is not a sensible method, and should not be used (Carpenter and Kenward, 2007). The LOCF is actually just an analysis of each patient's last observed value (Last Observation Analysis, LOA) and if LOA is really of interest then by definition the last observed measurement needs to be analyzed (Carpenter and Kenward, 2007).

Despite these criticisms of LOCF, if one wishes to "carry forward" information after withdrawal, then the appropriate distribution should be carried forward, not the observation (Carpenter and Kenward, 2007; Carpenter et al., 2013).

2.4.2.1.2 Baseline observation carried forward

The BOCF approach, on the other hand, replaces the missing values for each patient in the study with their respective baseline observations. This approach assumes that responses for dropouts remain the same as they were measured before randomization. This assumption is very intuitive in a clinical study involving HIV+ patients or patients with chronic cancer (Carpenter et al., 2013). However, this assumption might not always hold since most patients in a given study might have obtained significant benefit from the intervention after randomization. In such a case, patients may not necessarily loss all the benefit obtained after randomization.

2.4.3 Some remarks on the simple imputation methods

The LOCF and the BOCF methods are widely used in clinical trials. These methods entail restrictive assumptions that are unlikely to hold. Also, the uncertainty of imputation is not taken into account as the imputed values are identical to the observed values. These drawbacks in using the

LOCF and the BOCF may biased estimates of treatment effects and inflated rates of false-positive and false-negative results are likely (Molenberghs and Kenward, 2007; NRC, 2010; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Ayele et al., 2014). However, the use of simple methods set a historical example that, when combined with the desire to compare results with historical findings and the belief that LOCF and BOCF yielded "conservative" estimates of treatment effects, encouraged continued use of these methods (Mallinckrod et al., 2008). "Conservative" in this context was often interpreted as underestimating the magnitude of a treatment's benefit, thereby providing additional protection against erroneous approval of ineffective interventions (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013). Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf (2013) noted the potentially-appealing nature of these methods, but pointed out that conservatism is difficult to achieve in practice because underestimating a treatment effect, while conservative in a superiority test, could be anti-conservative in a non-inferiority test; this would certainly be anti-conservative for testing safety outcomes. A superiority trial is designed to detect a difference between treatments, whereas non-inferiority trials test whether a new experimental treatment is not unacceptably less efficacious than an active control treatment already in use. In addition, the anticipated conditions that would render an estimate conservative frequently are not realized in actual practice (Mallinckrod et al., 2008).

Fortunately, the problem of missing data has been an active area of investigation, and there have been many advances in statistical theory and in our ability to implement that theory. Furthermore, advances in statistical theory and in computer hardware and software have made many alternative methods simple and easy to implement. An alternative approach for handling missing data in a clinical trial setting is to use methods that are valid under the MAR assumption (Verbeke and Molenberghs, 2000; Little and Rubin, 2002; Molenberghs and Verbeke, 2005; Molenberghs and Kenward, 2007). Methods valid under MAR are also valid if data are MCAR. However, the reverse does not hold. Alternative approaches to the LOCF and BOCF limitations are the multiple imputation and the direct likelihood-based methods.

2.4.4 Multiple imputation

The multiple imputation (MI) approach assumes that the data are missing at random (MAR) and creates a set of K imputed values for the missing values there-by creating K-complete data sets (Rubin, 1976, 1996). The MI approach has the advantage of handling nonresponse in large data sets because it moves missing data burden from data analysis to data procedure. The MI approach also accounts for the uncertainty inherent in the imputation of the unobserved responses.

Several imputation methods proposed drawing missing values $\mathbf{Y}_i^m$ from the conditional distribu-

tion of unobserved responses $\mathbf{Y}_i^m$, given the observed responses $\mathbf{Y}_i^o$ and any potential covariate, $\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of parameter estimates describing the measurement process (Rubin, 1976, 1996). Imputation of missing values relies on the Bayesian posterior predictive distribution (Rubin, 1996; Daniels and Hogan, 2008).

Let $\hat{\boldsymbol{\theta}}$ be an estimator of the parameter $\boldsymbol{\theta}$ describing the complete data, with associated estimates $\hat{\boldsymbol{\Sigma}}$ of the variance-covariance $\boldsymbol{\Sigma}$ matrix. Multiple imputation draws missing values from the Bayesian posterior predictive distribution of the missing values under a specific Bayesian model for both $\mathbf{Y}_i^o$ and $\mathbf{Y}_i^m$. This means that the K-completed data analyses corresponding to K-imputations under one model give rise to K complete-data statistics. These statistics are then combined to single parameter estimates that appropriately adjust for the lost information due to imputation of missing data. Thus, the values of $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ calculated on the K- completed data sets are $\{\hat{\theta_1}, \ldots, \hat{\theta_k}\}$ and $\{\hat{\Sigma_1}, \ldots, \hat{\Sigma_k}\}$ respectively.

Rubin (1987) showed that multiple imputation relies on the Bayesian posterior distribution of the parameter $\boldsymbol{\theta}$, defined as

$$\Pr(\boldsymbol{\theta} \mid \mathbf{Y}_i^o) = \int \Pr(\boldsymbol{\theta} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m) \Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o) \, \mathrm{d}\mathbf{Y}^{\mathrm{m}}{}_i,$$

where $\Pr(\boldsymbol{\theta} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m)$ is the average over the repeated imputation from $\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)$. Two main estimators of interest are:

(a) the final estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, defined as

$$\mathrm{E}(\boldsymbol{\theta} \mid \mathbf{Y}_i^o) = \mathrm{E}[\mathrm{E}(\boldsymbol{\theta} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m) \mid \mathbf{Y}_i^o];$$

where $\mathrm{E}(\boldsymbol{\theta} \mid \mathbf{Y}_i^o)$ is the posterior mean $\boldsymbol{\theta}$ and $\mathrm{E}[\mathrm{E}(\boldsymbol{\theta} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m) \mid \mathbf{Y}_i^o]$ is the average of the complete data mean of $\boldsymbol{\theta}$ and

(b) the final estimator $\hat{\boldsymbol{\Sigma}}$ of the variance parameter $\boldsymbol{\Sigma}$, defined as

$$\mathrm{var}(\boldsymbol{\Sigma} \mid \mathbf{Y}_i^o) = \mathrm{E}[\mathrm{var}(\boldsymbol{\Sigma} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m) \mid \mathbf{Y}_i^o] + \mathrm{var}[\mathrm{E}(\boldsymbol{\Sigma} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m) \mid \mathbf{Y}_i^o],$$

where $\mathrm{var}(\boldsymbol{\Sigma} \mid \mathbf{Y}_i^o)$ is the posterior variance of $\boldsymbol{\Sigma}$, $\mathrm{E}[\mathrm{var}(\boldsymbol{\Sigma} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m) \mid \mathbf{Y}_i^o]$ is the average of the complete data variance $\boldsymbol{\Sigma}$ and $\mathrm{var}[\mathrm{E}(\boldsymbol{\Sigma} \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m) \mid \mathbf{Y}_i^o]$ is the variance of the repeated complete data posterior mean of $\boldsymbol{\theta}$.

These formulations are the basis for multiple imputation which are known as Rubin's MI rules.

The MI estimate of $\boldsymbol{\theta}$ using the Rubin's rule is the overall point estimate. Hence, the MI estimate is the average of the K estimates of $\boldsymbol{\theta}$ from the K imputed datasets defined by $\hat{\boldsymbol{\theta}}_K = \sum_{k=1}^{K} \frac{\boldsymbol{\theta}_l}{K}$. The variance-covariance of the estimates of the $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_K$ from K-imputed data sets is defined as $\boldsymbol{\delta}_K = \hat{\boldsymbol{\Sigma}}_K + \frac{K+1}{K}\mathbf{M}_K$, where $\hat{\boldsymbol{\Sigma}}_K = \sum_{k=1}^{K} \frac{\boldsymbol{\Sigma}_l}{K}$, $\hat{\mathbf{M}}_K = \sum_{k=1}^{K} \frac{(\boldsymbol{\theta}_l - \hat{\boldsymbol{\theta}})'(\boldsymbol{\theta}_l - \hat{\boldsymbol{\theta}})}{K-1}$, and $\hat{\mathbf{M}}_K$ is the estimator of

$\mathbf{M}_K$. The $\boldsymbol{\delta}_K$ estimates the variance of $\boldsymbol{\theta}$ given the $K-$completed data, $\hat{\boldsymbol{\Sigma}}_K$ is the average of the within-imputation covariance matrix if the data were complete, $\frac{K+1}{K}\mathbf{M}_K$ estimates the increase in variance due to missing data and $\mathbf{M}_K$ is the between-imputation covariance matrix.

The MI is attractive because it can be highly efficient, even for small values of K (Rubin, 1976, 1977, 1996). In some applications, merely 3 to 5 imputations are sufficient to obtain excellent results. An attraction of the Rubin's MI rules is that they lead to valid (if not statistically efficient) estimates conditional on using a finite number imputations.

2.4.5 Direct maximum likelihood analysis

The likelihood-based methods can be applied to incomplete longitudinal data (i.e., without any pre-processing or prior treatment of the missing values). These methods can also be applied to incomplete longitudinal data either after deletion (e.g., CC analysis) or after imputation of the missing observations (eg., MI under MAR). Under the CC analysis, since missing values are no longer present in the set of complete cases or in the imputed data set, likelihood approaches are based on the full-data likelihood (3.2) of the complete data. On the contrary, for incomplete longitudinal data, any method within a likelihood framework would require working with the observed-data likelihood (3.3). For MAR,

$$\ell_i = \ln \Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i^o, \boldsymbol{\psi}\right) + \ln \Pr\left(\mathbf{Y}_i^o \mid \boldsymbol{\theta}\right).$$

Maximum likelihood estimation would thus simply entail separate maximization of the component terms, which would translate, for instance, to fitting two maximum likelihood models: one for the observed responses and another for the non-responses conditional on the observed responses. Additional simplification arises for the case of MCAR,

$$\ell_i = \ln \Pr\left(\mathbf{R}_i \mid \boldsymbol{\psi}\right) + \ln \Pr\left(\mathbf{Y}_i^o \mid \boldsymbol{\theta}\right),$$

where the model for the non-response need not be conditional on the observed responses. Moreover, if the focus of inferences lies on the response process parameters $\boldsymbol{\theta}$, estimation of the (conditional) non-response model (given the observed measurements) can altogether be bypassed. As such, the direct likelihood approach is no more complicated than fitting a likelihood-based model on the complete cases. Standard software procedures that allow for incomplete observations, ensuring that the correct form of the likelihood is manipulated, would be able to obtain such a solution.

Direct likelihood under non-ignorability (e.g., MNAR) is a lot less straightforward in comparison with the ignorable case. Unlike the latter, the former does not admit further simplification of the observed-data log-likelihood contributions, due to the dependence of non-response on the unobserved

outcomes,

$$\ell_i = \ln L_i = \ln \int \Pr\left(\mathbf{Y}_i^o, \mathbf{Y}_i^m \mid \boldsymbol{\theta}\right) \Pr\left(\mathbf{R}_i \mid \mathbf{Y}^o, \mathbf{Y}^m, \boldsymbol{\psi}\right) \mathrm{dY}_i^m. \tag{2.1}$$

The integration over the missing values brings about additional levels of complexity to the direct likelihood approach for non-ignorable missingness. Also, evaluation or approximation of the integral to compute ℓ_i in (2.1), especially for high dimensions of missingness, can be computationally demanding (Molenberghs et al., 2008).

There is a little difference between the direct maximum likelihood (ML) and multiple imputation approaches under normality assumption for the observed data (Carpenter and Kenward, 2007). The likelihood-based methods can be considered as imputation methods regardless of whether missingness is ignored or modeled (Fitzmaurice et al., 2008, pp.401-403). When missing data is described as MAR and ignorable, likelihood-based methods are effectively imputing the missing values by modeling and estimating parameters for the joint distribution of the responses, $\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \boldsymbol{\theta}\right)$. When missing-data mechanism is described as ignorable, the likelihood-based methods impute missing values based on the marginal distribution of observed data $\Pr\left(\mathbf{Y}_i^o \mid \boldsymbol{\theta}\right)$ (Rubin, 1976). This means that the maximum likelihood estimates can be obtained by maximizing the likelihood function $\ell_i\left(\mathbf{Y}_i^o, \mathbf{X}_i, \boldsymbol{\theta}\right)$.

Rubin (1976) showed that likelihood-based inferences can be obtained by integrating over the missing responses from the joint distribution of the responses $\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \boldsymbol{\theta}\right)$, defined by

$$L\left(\boldsymbol{\theta}\right) \propto \prod_{i=1}^{N} \int \Pr\left(\mathbf{Y}_i^o, \mathbf{Y}_i^m \mid, \mathbf{X}_i, \boldsymbol{\theta}\right) \mathrm{dY}_i^m. \tag{2.2}$$

Intuitively, the missing values $\mathbf{Y}_i^m$ are validly predicted by the observed data via the model for conditional mean, $\mathrm{E}\left(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{X}_i, \boldsymbol{\theta}\right)$. This form of imputation becomes more transparent when expectation maximization (EM) (Dempster et al., 1977) algorithm is used.

The EM algorithm is an iterative procedure which allows us to compute the maximum likelihood (ML) estimates in the presence of missing data. Each iteration of the EM algorithm consists of two processes (steps). These steps alternate between (1) filling in the missing values with their conditional means, given the observed responses and parameter estimates from the previous iteration (expectation or E-step) and (2) maximization of the likelihood from the resulting "complete data" (maximization or M-step).

The EM algorithm is closely related to the following ad hoc process of handling missing data: (1) fill in the missing values by their estimated values, (2) estimate the parameters for this completed dataset, (3) use the estimated parameters to re-estimate the missing values, and (4) re-estimate the parameters from this updated completed dataset. Informally, it proceeds as

E-step : In the E-step, the missing data are estimated, given the observed data and current estimate of the model parameters. This is achieved by using the conditional expectation, $\mathrm{E}\left(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{X}_i, \boldsymbol{\theta}\right)$.

M-step : In the M-step, the likelihood function is maximized under the assumption that the missing data are known. This means that the estimate of the missing data from the E-step are used in place of the actual missing data.

These steps alternate until convergence of the parameter estimates is achieved. The EM algorithm is an iterative procedure for obtaining the ML estimate of $\boldsymbol{\theta}$ that maximizes the likelihood function 2.2. In the presence of missing data, the EM algorithm provides a natural framework for their inclusion. The algorithm achieves this by treating missing values as parameters after obtaining $\boldsymbol{\theta}$ from $\Pr\left(\mathbf{Y}_i^o \mid \boldsymbol{\theta}\right)$. As discussed under MAR in Section 2.2.2.2, MAR and MCAR are often referred to as ignorable mechanisms, ignorable in a sense that as long as one can establish that $\Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i, \mathbf{X}_i\right)$ is independent of the $\mathbf{Y}_i^m$, the $\Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i, \mathbf{X}_i\right)$ can be ignored and valid likelihood-based analysis can be obtained through a correctly specified joint distribution model for $\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i\right)$.

2.4.6 Comparison of the direct maximum likelihood and multiple imputation approaches

The direct maximum likelihood (ML) method and the multiple imputation (MI) methods are known to produce efficient estimates. However, one has an advantage over the other in some scenarios. The ML method is more efficient and produces correct standard errors compared with MI. Full efficiency for MI requires an infinite number of data sets. For a given data set, ML always gives the same results, whereas MI gives a different result each time it is used. However, one can "force" the MI to give the same results by setting the "seed". With MI, there is always potential conflict between the imputation model and the analysis model (Fitzmaurice et al., 2008). There is no conflict with ML because only one model is required. The ML do not require a model for the missing data mechanism. Rather, it "predicts" missing values implicitly by maximizing the likelihood function.

When the imputation model uses the same variables in the substantive model (assumed model for the measurement process), estimates from the ML and MI methods are comparable but the estimates from ML are more efficient. On the other hand, when the imputation model uses additional variable to improves its predictive power, or when particular forms of NMAR mechanism are relevant, the MI has an advantage over the ML method (Su et al., 2011). When the data are MAR, MI can lead to consistent, asymptotically efficient, and asymptotically normal estimates. The ML requires specialized software and it may be challenging and time-consuming. Once missing data are obtained, MI can be used with any kind of data and model with conventional software.

Although full data likelihood functions exist for marginal models (Molenberghs and Verbeke, 2006), however, under non-normal linear model setting, marginal models are computationally demanding since the likelihood function has no close form. If need be, this requirement can be avoided by specifying the likelihood only partially, resulting in a semi-parametric method (Bahadur, 1961; Molenberghs and Lesaffre, 1994; Liang and Zeger, 1986). Generally, the application of semi-parametric methods is not exclusively restricted to the area of longitudinal data, though such methods have gained popularity, particularly for the case of categorical (e.g., binary) repeated measures.

Under binary repeated measures, fully specified marginal models (Bahadur, 1961; Molenberghs and Lesaffre, 1994) exist and can be fitted; however, the intricacies can be restrictive (Sotto, 2009). As an alternative, Liang and Zeger (1986) proposed generalized estimating equations (GEE), which can be used to obtain marginal models for non-Gaussian longitudinal data, but, at the same time, avoiding the computational complexity of full likelihood. This approach does not rely on specification of a likelihood function for the repeated measurements but assumes a model for the means response and a model for variance-covariance structure. However, due to the non-likelihood nature of the GEEs, additional issues arise when data are missing. The issue with GEEs is that they are moment based estimators, and sample moments are biased when the data are MAR or NMAR. Under MAR, weighted GEE or the inverse probability weighting (IPW) (Seaman and White, 2013) provides consistent estimators (Robins et al., 1995; Robins and Rotnitzky, 1995; Scharfstein et al., 1999). The idea behind the weighted GEE is to weight each patient's contribution in the GEEs by the inverse probability that a subject drops out at the time he/she dropped out.

2.5 Standard methods of analysis

In this section, we describe the data structure and methods of analysis of such data. It is common to have grouped data in almost all areas of statistical application. The grouped data structure may be simple or complex depending on the grouping factor. A grouping factor could be treatment groups, age categories, or the sex of a patient. A simple structure occurs where each observation belongs to a single group and there is only one grouping factor. A complex structure has a hierarchical or nested structure. These data structures share a common characteristics of correlation of observation within the same group and between groups. Therefore, any method of analysis that assumes that observations are independent will produce invalid inference. These types of grouping structures are modeled appropriately by introducing random effects into the model (Searle, 1995; Laird and Ware, 1982; Verbeke and Molenberghs, 2009).

A model with both fixed and random effects terms is called a mixed-effect model. The mixed effects model is used primarily to describe a relationship between a response variable and covariate in the

data that are grouped according to one or more classification variables. Examples of such grouped data are typically obtained from longitudinal study or block designs. Observations within the same subject usually cannot be considered independent and mixed effects models provide a framework for modelling such data structure (Laird and Ware, 1982). The mixed effect models accounts for correlation between longitudinal measurements by introducing subject-specific random effect, which captures all unobserved subject-specific characteristics.

There are two commonly used mixed effect methods for analyzing longitudinal data. These methods are the linear mixed model (LMM) and the generalized linear mixed model (GLMM). These methods are similar in a sense that they account for variability between longitudinal measurement by introducing random effects. The only difference between them is that the LMM is used to model only normal responses whereas the GLMM can be used to model both normal and non-normal responses. We give a detailed review of the linear mixed effect model because the data used in this thesis are normal data.

2.5.1 Linear mixed effects model (LMM)

This section presents a review of the linear mixed effects model (LMM) (Laird and Ware, 1982). The LMM is widely used for repeated measures or longitudinal data where data are grouped. This model can be viewed as an extension of the classical linear model by introducing random effects. Detailed introduction to LMM can be found in (Verbeke and Molenberghs, 2009)

The following models are special cases of the LMM: (1) Variance components models (Searle et al., 2009), (2) mixed effects ANOVA models (Miller, 1977) and (3) linear models for longitudinal data (Laird and Ware, 1982). The origin of LMM for longitudinal or clustered data can be traced back to the ANOVA paradigm (Scheffe, 1999) and to the seminal paper by Harville (1977). The LMM usefulness for analyzing longitudinal data, especially in the life sciences, was made known by Laird and Ware (1982). The idea of allowing certain regression coefficients to vary randomly across individuals has been reported in the early contributions to growth curve analysis (Wishart, 1938; Box, 1950; Rao, 1958; Potthoff and Roy, 1964; Grizzle and Allen, 1969). These early contributions to growth curve modelling laid the foundation for the linear mixed-effect models.

The idea of allowing regression coefficients to vary across patients was also a common feature in the two-stage approach to analyzing longitudinal data. The two-stage formulations assumes that the repeated measurements on each individual follow a regression model with different regression coefficients for each individual. In the second stage of the two-stage formulation approach, distribution of these individual-specific regression parameters, or "random effects," is modeled. Rao (1965) proposed a formal framework of a two-stage approach by specifying a parametric growth curve model that assumed normally distributed random growth curve parameters. The approach is

simple and useful but the two-stage formulation of the linear mixed effects model becomes inconvenient to use because (1) in the first stage, the covariates were restricted to be time-varying with the exception of the intercept, (2) time-invariant covariates are only allowed in the second stage, where the individual-specific regression coefficients were modeled as a linear function of these covariates and (3) placed unnecessary, and often inconvenient, constraints on the choice of the design matrix for the fixed effects.

Nonetheless, the two-stage formulation provided motivation for the main ideas and concepts underlying linear mixed-effects models. Following the general class of mixed models introduced earlier by Harville (1977), Laird and Ware (1982) proposed a flexible class of linear mixed-effects models for longitudinal data. Laird and Ware (1982) formulation of the LMM had two desirable features. There were fewer restrictions on the design matrices for the fixed and random effects and the model parameters could be estimated efficiently via likelihood- based methods.

2.5.1.1 The form of the linear mixed effect model

The LMM is often used to analyzed repeated measures data by adding a subject-specific random effect to the model that captures all the unobserved subject-specific characteristics. Let $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{in})$ denotes an N-dimensional vector of the responses for the ith patient and $\mathbf{X}_i$ be an $N \times p$ design matrix of covariates for the ith patient. The linear mixed effect model (LMM) (Laird and Ware, 1982) is assumed for the measurement process and is given by

$$\begin{cases} \mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \\ \boldsymbol{b}_i \sim N\left(\mathbf{0}, \mathbf{G}(\boldsymbol{\rho})\right), \\ \boldsymbol{\epsilon}_i \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\sigma})\right), \\ \boldsymbol{b}_i \perp\!\!\!\perp \boldsymbol{\epsilon}_i, \end{cases} \tag{2.3}$$

where $\boldsymbol{b}_i$ is an q-dimensional vector of random effects, $\mathbf{Z}_i$ and $\mathbf{X}_i$ are $N \times q$ and $N \times q$ dimensional matrices of known covariates, $\boldsymbol{\beta}$ is a p-dimensional vector containing the fixed effects, $\boldsymbol{\epsilon}_i$ is an N-dimensional vector of residual components, $\mathbf{G}(\boldsymbol{\rho})$ and $\boldsymbol{\Sigma}(\boldsymbol{\sigma})$ are $q \times q$ and $n \times n$ covariance matrices respectively and $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ are $c \times 1$ and $s \times 1$ (with $s \leq n(n+1)/2$) vectors of unknown variance parameters corresponding to $\boldsymbol{\epsilon}_i$ and $\boldsymbol{b}_i$ respectively.

Widespread application of the LMM was hindered by difficulties of estimation of mixed effects models until Laird and Ware (1982) showed how the EM algorithm (Dempster et al., 1977) could be used to fit this general class of models for longitudinal data. Immediately after Laird and Ware (1982) paper, Jennrich and Schluchter (1986) proposed a variety of alternative algorithms such as Fisher scoring and Newton-Raphson. The EM algorithm (Dempster et al., 1977) or Newton-Raphson methods (Thisted, 1988) are used for parameter estimation in the mixed effect models.

However, Lindstrom and Bates (1988) noted that the Newton-Raphson methods seems to be more efficient than the EM algorithm.

There are several methods of parameter estimation for the LMM. The most common and widely used parameter estimation methods for the LMM are the maximum likelihood (ML) and the restricted maximum likelihood (REML) (Searle, 1995; Laird and Ware, 1982; Harville, 1977; Henderson, 1953). Detailed information on parameter estimation formulations, based on the ML and REML, can be found in Gumedze and Dunne (2011); Harville (1974); West et al. (2014); Henderson (1953); Henderson et al. (1959); Searle et al. (2009); Searle (2012); Patterson and Thompson (1971).

2.6 Description of the IMPI trial data

In this thesis, we used data from the IMPI trial (Mayosi et al., 2012, 2014). The IMPI trial was a multicentre international randomized doubled-blind placebo-controlled 2×2 factorial study. The IMPI trial tested prednisolone and Mycobacterium indicus pranii (M. indicus pranii) immunotherapy treatments in TB pericarditis (TBP) patients in Africa. There are four treatment arms. Patients were randomized to received combination of either M^+P^+ or M^+P^- or M^-P^+ or M^-P^-, where M^+ and M^- denote the M. indicus pranii and M. indicus pranii placebo respectively and P^+ and P^- denote prednisolone and prednisolone placebo respectively. TBP leads to high mortality, especially in countries with limited resources and with concomitant epidemics of human immunodeficiency virus (HIV) infection (Mayosi et al., 2012, 2014).

Tuberculous pericarditis is associated with high morbidity and mortality even if anti-tuberculosis treatment is taken as directed (Mayosi et al., 2014). A reduction in the strength of the inflammatory response in TB pericarditis may improve patients' conditions by reducing cardiac tamponade and pericardial constriction. However, whether the use of adjunctive immunomodulation with corticosteroids and M. indicus pranii can safely reduce mortality and morbidity is uncertain (Mayosi et al., 2014). To investigate whether adjunctive immunomodulation with corticosteroids and M. indicus pranii can safely reduce mortality and morbidity, Mayosi and colleagues set up the IMPI trial (Mayosi et al., 2012, 2014).

In total, 1400 patients with definite probable tuberculosis pericardial effusion, from 9 African countries in 19 centres were enrolled in the four-year trial. Eligible patients were randomly assigned to receive oral pill prednisolone for 6 weeks and M. indicus pranii or placebo for 3 months. Patients were followed up at weeks 2, 4, and 6 and months 3 and 6 during the intervention period and 6-monthly thereafter for up to 4 years (Mayosi et al., 2012).

The IMPI trial was conducted from January 2009 through February 2014 at 19 hospitals in eight African countries (Mayosi et al., 2014). For the comparison of prednisolone with placebo, 706

patients were assigned to receive prednisolone and 694 to receive placebo. For the comparison of M. indicus pranii with placebo, 625 were assigned to receive M. indicus pranii and 625 to receive placebo. This comparison was stopped early (on February 14, 2013) for futility (Mayosi et al., 2014). The trial was powered for a rate of non-adherence of 10% in the active-treatment groups. This rate was almost achieved (with non-adherence rate of 11%) in the prednisolone group and non-adherence rate was higher in the M. indicus pranii group (21%), owing mainly to injection-site side effects (Mayosi et al., 2014). Also approximately 18% deaths occurred due to TBP related (4.2%), TB related other than pericardial (3.3%), HIV related (1.3%), cardiovascular related (1%), unknown (3.5%)

The main aim of the IMPI trial was to assess the effectiveness and safety of oral pill prednisolone and M.w injection in reducing the time to first occurrence of the primary composite outcome of death, pericardial constriction, or cardiac tamponade requiring pericardial drainage in patients with TB pericardial effusion (Mayosi et al., 2014). In this thesis, we assessed the effect of trial medication (prednisolone) on CD4 count changes over time. The proportion of HIV and TB co-infected patients is 67%. Hence the interest in investigating the effect of prednisolone among HIV positive (denoted as HIV+) patients.

We restricted our analysis to HIV positive patients only who have at least two CD4 count values observed. In the IMPI trial, patients who were confirmed HIV+ at the time of randomization or confirmed to be HIV+ during the trial, were given a standard of care (ART) and their CD4 count levels were measured at some visits. Mayosi et al. (2014) results showed that the oral pill prednisolone and M. indicus pranii do not interact and hence, treatments arms were analysed separately with their corresponding placebo arms. Also, their results showed that prednisolone reduces the risk of constriction, whereas M. indicus pranii was not effective. We considered analysis of the CD4 count measurements under the prednisolone treatment and its corresponding placebo arm only.

The analysis of CD4 count data is restricted to the mandated periods for CD4 count measurements; baseline, week 2, months 1, 3 and 6. However, most South Africa centres continued to measure CD4 count at months 24, 36 and 48 scheduled visit time. These data were excluded in this analysis. Data on months 24, 36 and 48 were excluded from the analyses because, most South Africa centres continued to measure CD4 count at months 24, 36 and 48 scheduled visit time. Centres in other countries do not measure CD4 count at these visit. So inclusion of data from these time points into the analysis is likely to produce biased statistical inferences. A majority of patients had unobserved CD4 count with 72%, 84%, 93% as missingness proportions for the months 24, 36 and 48, respectively.

In this thesis, we applied the approaches to non-monotone and monotone missing CD4 count data.

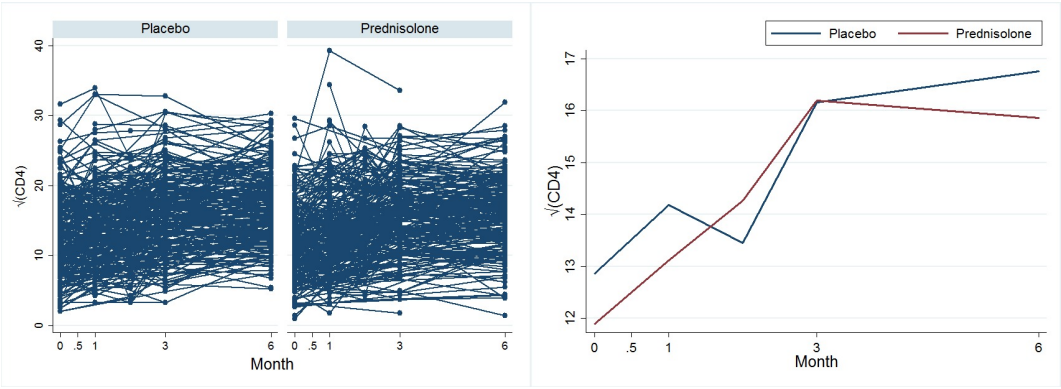

Figure 2.2: Profiles plots of the non-monotone $\sqrt{\text{CD4}}$ count data (left panel) and the mean $\sqrt{\text{CD4}}$ count (right panel), by placebo and prednisolone treatment arms

For non-monotone missing data pattern, patients can be missing at any scheduled visit and then be observed at the subsequent visit. For monotone missing data pattern, if the ith patient is missing at schedule visit j, then this same patient will be missing at the next scheduled visit $j + 1$.

The key scientific question to address is whether prednisolone interacts with ART. If it does, we expect a decline or an increase in the level of CD4 count measurements among patients in the prednisolone arm relative to that of the patients in placebo arm.

2.6.1 Non-monotone data

We consider HIV+ patients with at least two responses. Out of the 1400 patients enrolled in the IMPI trial, 587 were HIV+ with 294 patients in the placebo arm and 293 in the prednisolone arm. Some of the patients have missing values within the selected scheduled visits. The left panel of Figure 2.2 shows profiles plots of the observed $\sqrt{\text{CD4}}$ count measurements for each patient. Some of the patients CD4 count values are missing at either months 0.5, 1, or 3, after the baseline measurements are taken whereas some patients completed the study with their values observed from baseline up to month 6. Because there are too many patients in the left panel of the Figure 2.2, the figure is not very informative. We have provided observed profiles plots of 29 (5%) patients in Figure 2.3 to make this panel more informative. It can be observed from these plots that some patients completed the study (observed from baseline 0 to the month 6) while others have missing values (incomplete cases). The right panel of the Figure 2.2 shows the profiles plots of the mean $\sqrt{\text{CD4}}$ count measurements by treatment arms. The mean profile plots showed a slight reduction (after 3 months) of CD4 count level among patients in the prednisolone arm compared with those in the placebo arm. There are 25 missingness patterns, presented in Table 2.1. A missingness pattern represents time points for which a group of patients' values are missing or observed at all time points. The Table 2.1 shows the mean $\sqrt{\text{CD4}}$ count for each of the missingness patterns at each visit by treatment arm. The proportion of patients with missing values, in the prednisolone

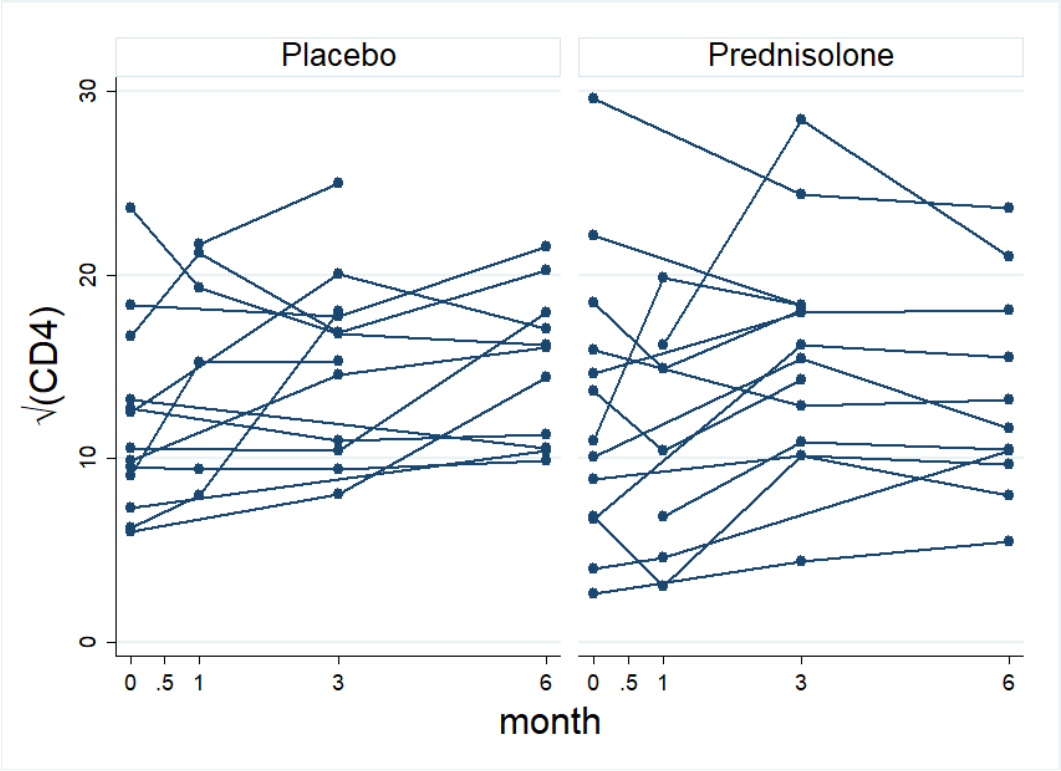

Figure 2.3: Observed profiles plots of the non-monotone $\sqrt{\text{CD4}}$ count data for 29 (5%) patients, by placebo and prednisolone treatment arms

arm (84%), is approximately the same as that of the patients in the placebo arm (85%). Table 2.1 presents summaries of the $\sqrt{\text{CD4}}$ count data by treatment groups. The distribution of the pattern of missingness between the two treatment groups does not differ (chi-squared test statistic = 29.97, p = 0.1858).

2.6.2 Monotone data

Out of the 587 patients, the monotone data consist of 137 HIV+ patients with 64 were in the placebo arm and 73 in the prednisolone arm. The left panel of Figure 2.4 shows the observed $\sqrt{\text{CD4}}$ count profiles plots for each HIV-positive subject. There was dropout at months 0.5, 1, and 3, whereas some subjects completed the study. The right panel of Figure 2.4 displays the mean $\sqrt{\text{CD4}}$ counts profiles plots by treatment arm and shows that there is a slight reduction of CD4 count level among those in the prednisolone arm compared with those in the placebo arm, and a linear trend in the CD4 count measurement in both the prednisolone and the placebo arms. Our initial analysis showed that the quadratic time component was not statistically significant, so the fitted linear mixed model included the linear trend component. Table 2.2 gives the number and proportion of patients remaining at each visit by treatment arms. There is a higher completion rate of 44 (69%) in the placebo arm, compared with 46 (63%) in the prednisolone treatment arm.

We have provided summaries of the main outcome ($\sqrt{\text{CD4}}$ count) stratified per dropout patterns and treatment groups, as a function of time in Table 2.3. There are four dropout patterns and the

Table 2.1: Non-monotone data: mean $\sqrt{\text{CD4}}$ count at each visit by treatment arm

	Visit (month)					
Missingness pattern[a]	**0**	**0.5**	**1**	**3**	**6**	N (%)
Prednisolone arm						
1	13.18	15.52	14.70	16.76	16.822	46 (16)
2	19.89	19.27	19.51	21.17	-	5 (2)
3	9.04	12.26	12.49	-	-	12 (4.1)
4	11.64	11.43	-	-	-	10 (3.4)
5	2.83	3.61	3.32	-	4.47	1 (0.3)
6	9.62	-	-	-	14.15	11 (4)
7	13.84	16.89	-	18.47	18.12	4 (1.4)
8	11.06	-	-	14.57	14.29	28 (10)
9	-	-	-	19.10	19.57	5 (2)
10	12.42	-	13.03	15.83	16.09	76 (26)
11	-	-	11.85	16.16	15.82755	13 (4.4)
12	-	12.27	13.99	15.88	15.66	14 (5)
13	9.07	-	9.26	-	15.92	12 (4.1)
14	-	-	-	-	-	0 (0)
15	12.52	-	13.21	15.54	-	14 (5)
16	14.12	-	-	15.73	-	9 (3.1)
17	8.12	-	10.74	-	-	13 (4.4)
18	-	14.35	14.73	16.00	-	2 (0.7)
19	-	17.10	-	20.95	21.59	4 (1.4)
20	-	15.07	13.10	-	-	3 (1)
21	-	10.21	-	-	10.80	2 (0.7)
22	-	-	15.07	16.78	-	3 (1)
23	-	-	13.09	-	9.32	3 (1)
24	-	18.19	-	14.04	-	1 (0.3)
25	-	8.35	8.20	-	13.36	2 (0.7)
All patients mean (std)	11.89 (0.333)	106, 14.27 (0.575)	219, 13.12 (0.397)	224, 16.20 (0.344)	221, 15.86 (0.348)	293 (100)

[a]*Missingness patterns: 1 = CD4 count data at all visits, 2 = CD4 count data at all visits except 6, 3 = CD4 count data at all visits except 6 and 3, 4 = CD4 count data at all visits except 6, 3, and 1, 5 = CD4 count data at all visits except visit 3, 6 = CD4 count data at baseline and visit 6 etc.*

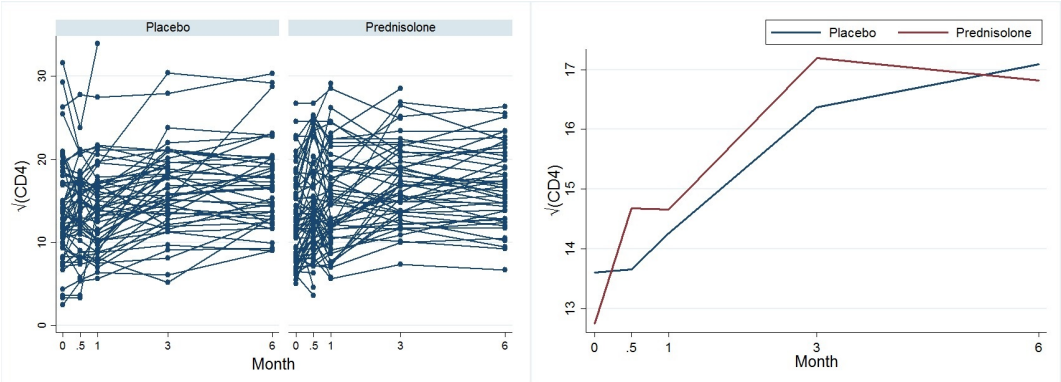

Figure 2.4: Individual profiles plots of the monotone $\sqrt{\text{CD4}}$ count data (left panel) and the mean $\sqrt{\text{CD4}}$ count (right panel), by placebo and prednisolone treatment arms

Table 2.2: Percentage of patients remaining in the study at each visit

	Placebo	Prednisolone
Month	N (%)	N (%)
0	64 (100)	73 (100)
0.5	64 (100)	73 (100)
1	57 (88)	63 (86)
3	53 (83)	51 (70)
6	44 (69)	46 (63)

Table 2.3 shows the mean $\sqrt{\text{CD4}}$ count for each of the patterns at each visit by treatment arm. The dropout patterns 4, 3, 2 and 1 represent completers (those patients who completed the study) and those who dropped out at months 3, 1 and 0.5 respectively. The distributions of the patterns of missingness between the two treatment groups do not differ (chi-squared test statistic = 5.15, $p = 0.161$).

Table 2.3: Monotone data: mean $\sqrt{\text{CD4}}$ counts at each visit by dropout pattern and treatment arm

	Dropout time (months)					
Dropout pattern[a]	0	0.5	1	3	6	N (%)
Placebo arm						
4	13.14	13.47	13.62	16.24	17.09	44 (69)
3	12.58	13.702	14.76	17.01	-	9 (14)
2	16.90	17.68	20.27	-	-	4 (6)
1	15.98	12.44	-	-	-	7 (11)
All patients mean (std)	13.61 (5.84)	13.65 (4.97)	14.26 (5.32)	16.37 (4.85)	17.09 (5.14)	64 (100)
Prednisolone arm						
4	13.18	15.52	14.70	16.76	16.82	46 (63)
3	19.89	19.27	19.51	21.17	-	5 (7)
2	9.04	12.26	12.49	-	-	12 (16)
1	11.64	11.43	-	-	-	10 (14)
All patients mean (std)	12.75 (5.10)	14.68 (5.71)	14.66 (5.84)	17.19 (4.80)	16.82 (4.72)	73 (100)

[a] *Dropout patterns: 4 = subjects who had all measurements up to 6 months (completers), 3 = subjects who had measurements up to 3 months, 2 = subjects who had measurements up to 1 month, and 1 = subjects who had measurements up to 2 weeks.*

Figure 2.5 shows the profile plots of the mean $\sqrt{\text{CD4}}$ count of the four deviation patterns for patients in the placebo and prednisolone groups. This figure gives an indication that the $\sqrt{\text{CD4}}$ count increases over time. Figure 2.5 agrees with the mean profiles plot in Figure 2.2 and Figure 2.4 that there is slight increase in the $\sqrt{\text{CD4}}$ count among patients in the placebo arm compare with those in prednisolone arm.

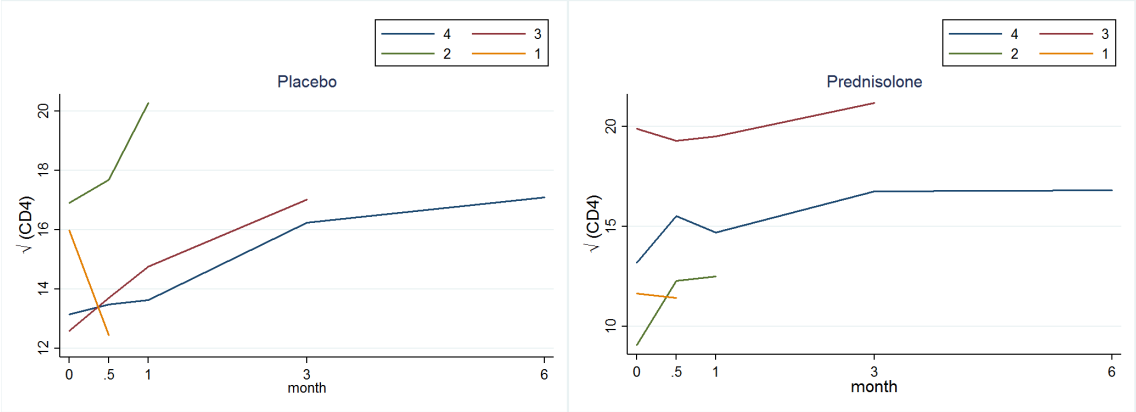

Figure 2.5: Monotone data: profile plots of the mean of the $\sqrt{\text{CD4}}$ count for each deviation pattern under the placebo arm (left panel) and the active arm (right panel). Blue pattern: group of patients who completed the study (completers); brown pattern: group of patients who dropped out after month 3; green pattern: group of patients who dropped out after month 1; and yellow pattern: group of patients who dropped out after week 2.

2.7 Summary and discussion

Since longitudinal data are correlated, there is the need to account for correlation between the repeated measurements in order to produce valid statistical inferences. In the data description Section 2.6, it can be observed the Figures 2.4 and 2.5 and Tables 2.1, 2.2, and 2.3 showed that the longitudinal CD4 count data are incomplete. Hence, method of analysis for incomplete longitudinal data must account for the missing data in the statistical analysis in order to yield valid statistical inferences. This requires the analyst to have an idea of the three missing data mechanisms classified by Rubin (Rubin, 1976) as missing completely at random (MCAR), missing at random (MAR), and not missing at random (NMAR).

A Method of analysis that yields valid statistical inferences in the absence of missing data will also produce valid statistical inferences when applied to missing data assumed to be MCAR. However, such method will not yield valid statistical inferences under MAR and NMAR mechanisms. The likelihood-based and the multiple imputation approaches can be applied to produce valid statistical inferences under MCAR and MAR mechanisms but not when data are NMAR. Under NMAR mechanism, the observed data do not provide information that can be used to either prove or disprove one NMAR mechanism over another. This means that testing for such assumption depends on untestable assumption and emphasis is placed on NMAR methods (presented in Chapter 3) for conducting sensitivity analysis to ascertain the type of missing data mechanism operating in a given trial. Sensitivity analysis should be considered by default when handling data with missing values since reader of a clinical report will place confidence in results supported with sensitivity analysis when there are missing data. Sensitivity analysis is main focus of this thesis.

Sensitivity analysis requires the analyst to carefully select primary analysis model for the trial at hand and then consider methods that stress-test the results of the primary analysis by making plausible alternative scientific assumptions that deviate from the primary analysis assumption. This means that selecting appropriate primary analysis is key to sensitivity analysis because a wrong primary analysis formulation will ultimately leads to wrong sensitivity analysis formulations. Primary analysis formulation depends on the whether the objective of the trial is on efficacy or effectiveness estimand (Carpenter et al., 2013; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013).

Based on our discussion in Section 2.3.1 and given the data description in Section 2.6, the estimand of interest in this thesis is effectiveness hypothesis since the trial aim is on response improvement at the plan end period for all randomized patients. This also suggests that the primary analysis under MAR mechanism would be appropriate since such assumption allows the use of post-deviation data either via imputation or if recorded after deviation/dropout. We then stress-test statistical inferences under this primary analysis model to assess their sensitivity to various alternative assumptions under Chapters 4-6. Since the $\sqrt{\text{CD4}}$ count measurements is normally distributed, we assumed the linear mixed effect model (Laird and Ware, 1982) for the observed data.

Chapter 3

Joint models for non-ignorable incomplete longitudinal data

3.1 Introduction

As we have stated in Chapter 1, to produce valid statistical inferences, we require a joint model for the measurement process and the missing data mechanism (see Molenberghs and Fitzmaurice in Fitzmaurice et al. 2008, Ch. 17, pp. 401-403). Hence the purpose of this chapter is to review the common joint models for non-ignorable missing data mechanism (NMAR). Missing data mechanism is said to be non-ignorable when the probability that a response will be missing still depends on the responses, given the observed data. These models are described as non-ignorable because the missing data mechanism cannot be ignored in the modeling process. The discussion in this chapter is necessary because sensitivity analysis methods in this thesis are based on these joint models for non-ignorable missing data. We discuss the National Research Council (NRC) panel view on the need to conduct sensitivity analysis to dropout and how to conduct such analysis. We discuss specifications in these joint models for non-ignorable missing data that allow us to perform various forms of sensitivity analysis presented in Chapters 4-6.

Little (1993) and Rubin (1976) identified two main classes of joint models for non-ignorable missing data mechanism. These classes of models are the selection and the pattern-mixture models (Wu and Bailey, 1988). The selection models use a complete-data model for the longitudinal outcomes, and then the probability of non-response is modeled conditional on the possibly unobserved outcomes. The selection models by specification are under-identified. This is because some of the parameter estimates in the selection model depend on unobserved outcome. That is, with selection models, identification comes from unverifiable models for the dependence of the non-response probabilities on the unobserved outcomes.

The pattern-mixture models use a model for the conditional distribution of the outcomes given non-response patterns and then a model for non-response. The distribution of the outcome, given patterns of non-response, is not completely identifiable. This is because among the incomplete patterns, certain response variables are not observed. Hence, restrictions must be built into the model to ensure that there are links among the distributions of the outcomes conditional on the

patterns of non-response (Little, 1993; Wang-Clow et al., 1995; Little, 1995; Molenberghs et al., 2008).

Some of the earlier work on methods for handling non-ignorable missing data in longitudinal studies was conducted by Wu and Bailey (1988); Wu and Carroll (1988). They proposed the selection modeling approach for non-ignorable dropouts. Their selection modeling approach assumes that the continuous responses follow a simple linear random-effects model and that the dropout process depends upon an individual's random intercept and slope. The models where the dropout probabilities depend indirectly on the distribution of the unobserved responses, via the random effects, are often referred to as shared-parameter models (Fitzmaurice et al., 2008; Molenberghs et al., 2008; Creemers et al., 2010, 2011).

An alternative selection modelling approach proposed by Diggle and Kenward (1994) allows the probability of non-response to depend directly on the unobserved outcome rather than on underlying random effects.

3.1.1 The selection modelling framework

The selection model can be viewed as a multivariate model where one variable represents the marginal density of the measurements process and the other variable represents the conditional density of the missingness process, given the outcomes.

Some of the earlier work on methods for handling non-ignorable missing data in longitudinal studies was conducted by Wu and colleagues (Wu and Carroll, 1988). Wu and Carroll (1988) proposed a selection modelling approach, which is now being used by many subsequent researchers. Their selection modelling approach assumes that the continuous responses follow a simple linear random-effects model (Laird and Ware, 1982) and that the dropout process depends upon an individual's random intercept and slope. Models where the dropout probabilities depend indirectly upon the distribution of the unobserved responses, via the random effects, are often referred to as shared-parameter models (Fitzmaurice et al., 2008, Ch. 13 and Ch. 19). We discuss shared-parameter models in Section 3.1.3. An alternative selection modelling approach, based on earlier work in the univariate setting by econometricians Heckman (Heckman, 1976), was proposed by Diggle and Kenward (1994). These author allowed the probability of non-response to depend directly on the distribution of the unobserved outcome rather than on underlying random effects.

Heckman (1976) introduced a Tobit model for a potentially missing continuous outcome such that this model combines a marginal Gaussian regression model for the measurement process with a probit model for the missingness process. However, the structure of this model and its associated estimation approach was subjected to much debate. This debate focus was on the problem of

identifiability (data not enough to estimate all parameters specified by the model) and sensitivity (robustness of inferences to changing assumptions about the missing data). This basic structure of the Tobit model forms the basis of the simplest forms of the selection models that have been proposed in a wide range of disciplines (Molenberghs et al., 2014, pp.55). In these models, a suitable model such as a multivariate normal model for the measurement process is combined with a binary regression model for the missingness process. Often, logit of the probability of dropout at each time point is regressed on the previous and current outcomes.

The selection model (SeM) can be defined as (Rubin, 1976; Little and Yau, 1996)

$$\Pr\left(Y_{ij}, D_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(Y_{ij} \mid \boldsymbol{\theta}\right) \Pr\left(D_i \mid Y_{ij}, \boldsymbol{\psi}\right), \tag{3.1}$$

where $\boldsymbol{\theta}$ is a vector of parameters describing the measurement process, $\boldsymbol{\psi}$ is a vector of parameters describing the missing data mechanism and D_i indicates a scheduled visit time where the ith patient dropped out. The first component $\Pr\left(\mathbf{Y}_{ij} \mid \boldsymbol{\theta}\right)$ of this selection model (3.1) is the marginal density of the measurement processes and the second component $\Pr\left(D_i \mid \mathbf{Y}_{ij}, \boldsymbol{\psi}\right)$ is the conditional density of missingness and the measurement processes. Rubin (1976) showed that when the missing data mechanism is described (MAR) and $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ are functionally independent, the likelihood-based and Bayesian methods yield valid inferences (Verbeke and Molenberghs, 2009; Fitzmaurice et al., 2008).

The likelihood-based methods involve maximization of the full-data likelihood function

$$L^* \equiv L^*\left(\boldsymbol{\theta}, \boldsymbol{\psi} \mid \mathbf{Y}, D\right) = \prod_{i=1}^{N} \Pr\left(\mathbf{Y}_i, D_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \prod_{i=1}^{N} \Pr\left(\mathbf{Y}_i^o, \mathbf{Y}_i^m, D_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right). \tag{3.2}$$

In the presence of missing values, however, inferences must be based on what is observed, and thus, the full-data likelihood L^* must be replaced by the observed-data likelihood L, for which the individual likelihood contributions need to be integrated over the missing values defined as

$$L \equiv L\left(\boldsymbol{\theta}, \boldsymbol{\psi} \mid \mathbf{Y}, \mathbf{D}\right) = \prod_{i=1}^{N} \int \Pr\left(\mathbf{Y}_i^o, \mathbf{Y}_i^m, D_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right) \mathrm{dY}_i^m. \tag{3.3}$$

Restricting our focus on the observed-data likelihood, the contribution of the i^{th} subject, under the SeM framework (3.1), the integral in (3.3) becomes

$$L_i \equiv L_i\left(\boldsymbol{\theta}, \boldsymbol{\psi} \mid \mathbf{Y}_i, D_i\right) = \prod_{i=1}^{N} \int \Pr\left(\mathbf{Y}_i^o, \mathbf{Y}_i^m \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right) \Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m \mid \boldsymbol{\psi}\right) \mathrm{dY}_i^m. \tag{3.4}$$

Rubin (1976) classified the missing data mechanism according to the selection model (3.1) as follows. For MCAR, the missing data mechanism component is defined as $\Pr\left(D_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(D_i \mid \boldsymbol{\psi}\right)$. It follows that the selection model can now be written as

$$\Pr\left(\mathbf{Y}_i^o, D_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(\mathbf{Y}_i^o \mid \theta\right) \Pr\left(D_i \mid \boldsymbol{\psi}\right). \tag{3.5}$$

The implication of model (3.5) is that we can now model the measurement and the missing data mechanism separately. This further implies that standard methods of analysis can be applied to

produce valid inference. This is because, the dropout D_i process is independent of the observed $\mathbf{Y}_i^o$ and the unobserved $\mathbf{Y}_i^o$ components of $\mathbf{Y}_i$.

For MAR, the missing data mechanism component is defined as $\Pr(D_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\theta}, \boldsymbol{\psi}) = \Pr(D_i \mid \mathbf{Y}_i^o, \boldsymbol{\psi})$ and the selection model can now be written as

$$\Pr(\mathbf{Y}_i^o, D_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}) = \Pr(\mathbf{Y}_i^o \mid \boldsymbol{\theta}) \Pr(D_i \mid \mathbf{Y}_i^o, \boldsymbol{\psi}). \tag{3.6}$$

The implication of model (3.6) is that valid inferences cannot be obtained without accounting for the missing data mechanism. This is because, the dropout D_i process depends on the observed $\mathbf{Y}_i^o$ component of $\mathbf{Y}_i$. However, under the ignorable condition, the observed data and the missing data models can be fitted separately to obtain valid inferences.

For NMAR, the missing data mechanism component is defined as $\Pr(D_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\theta}, \boldsymbol{\psi}) = \Pr(D_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m \boldsymbol{\psi})$. This defines the non-ignorable missingness, and the selection model can now be written as

$$\Pr(\mathbf{Y}_i^o, D_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}) = \int \Pr(\mathbf{Y}_i^o, \mathbf{Y}_i^m \mid \boldsymbol{\theta}) \Pr(D_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\psi}) \, \mathrm{d}\mathbf{Y}_i^m. \tag{3.7}$$

The implication of model (3.7) is that valid inference cannot be obtained without accounting for the missing data mechanism. This is because the dropout D_i process depends on the observed $\mathbf{Y}_i^o$ and the unobserved $\mathbf{Y}_i^m$ components of $\mathbf{Y}_i$. The model (3.7) is a joint model for non-ignorable missing data. This is because the measurement and the missingness processes have to be modeled jointly to obtained valid inferences.

Rubin (1976) showed that when data are assumed to be MAR and $\boldsymbol{\psi}$ and $\boldsymbol{\theta}$ are functionally independent, valid inference can be obtained by modeling the measurement and the missingness processes separately. This is called ignorability. This means that the likelihood-based inference remains valid when the missing data mechanism is ignored. In this case, the likelihood of interest is then based upon the observed data.

The i^{th} patient contribution to the likelihood based on (3.6) is of the form

$$L_i^*(\boldsymbol{\theta}, \boldsymbol{\psi} \mid \mathbf{Y}_i, \mathbf{R}_i) \propto (\mathbf{Y}_i, \mathbf{R}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}). \tag{3.8}$$

Since inference has to be based on what is observed, the full data likelihood L_i has to be replaced by the observed data likelihood defined as L_i^o

$$L_i(\boldsymbol{\theta}, \boldsymbol{\psi} \mid \mathbf{Y}_i^o, \mathbf{R}_i) \propto (\mathbf{Y}_i^o, \mathbf{R}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}) \tag{3.9}$$

with

$$\begin{aligned} \Pr(\mathbf{Y}_i^o, \mathbf{R}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}) &= \int \Pr(\mathbf{Y}_i, \mathbf{R}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}) \, \mathrm{d}\mathbf{Y}_i^m \\ &= \int \Pr(\mathbf{Y}_i^o, \mathbf{Y}_i^m \mid \boldsymbol{\theta}) \Pr(\mathbf{R}_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\psi}) \, \mathrm{d}\mathbf{Y}_i^m \end{aligned} \tag{3.10}$$

Under MAR mechanism, the model (3.10) can be written as

$$\begin{aligned}\Pr\left(\mathbf{Y}_i^o, \mathbf{R}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right) &= \int \Pr\left(\mathbf{Y}_i^o, \mathbf{Y}_i^m \mid \boldsymbol{\theta}\right) \Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i^o, \boldsymbol{\psi}\right) \mathrm{d}\mathbf{Y}_i^m \\ &= \Pr\left(\mathbf{Y}_i^o \mid \boldsymbol{\theta}\right) \Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i^o, \boldsymbol{\psi}\right)\end{aligned} \tag{3.11}$$

Since $\mathbf{Y}_i$ and $\mathbf{R}_i$ are independent, valid likelihood-based inference can obtained by fitting $\Pr\left(\mathbf{Y}_i^o \mid \boldsymbol{\theta}\right)$ and $\Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i^o, \boldsymbol{\psi}\right)$ separately (Rubin, 1976, 1996). For Bayesian inferences, a prior distribution is assumed for each of the parameters in the selection model (Daniels and Hogan, 2008, pp. 168), and valid inferences can be obtained if besides the separability condition (when $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ are distinct), the priors are independent (Little and Rubin, 1987; Daniels and Hogan, 2008, pp. 168). A posterior distribution is then constructed (Daniels and Hogan, 2008, see pp. 180-181), using the data likelihood and the prior distributions of the parameters, to sample/obtain parameter estimates in the selection model. Complexity often arise in the integration of likelihood with respect to the missing response and Bayesian MCMC approaches avoid direct evaluation of this integral through data augmentation (Daniels and Hogan, 2008, pp. 178).

We note that when the separability condition is satisfied, within the likelihood framework, ignorability is equivalent to the MAR or MCAR. This implies that non-ignorability and MNAR are equivalent in this context. A formal derivation given by Rubin (1976), and Little and Rubin (1987), showed that the same requirements hold for Bayesian inferences. The implication of the ignorability is that a software module with likelihood estimation facilities and with the ability to handle incompletely-observed subjects, manipulates the correct likelihood and thus provides valid parameter estimates, standard errors if based on the observed information matrix, and likelihood ratio values (Molenberghs et al., 1998). This result makes direct-likelihood analyses viable candidates for the status of primary analysis in clinical trials (Molenberghs et al., 2004).

3.1.2 The pattern-mixture modeling framework

The pattern-mixture model (PMM) framework is a reverse factorization of the selection model (Heckman, 1976; Rubin, 1976; Diggle and Kenward, 1994; Little and Yau, 1996). The PMM approach can be specified as

$$\begin{aligned}\Pr\left(\mathbf{Y}_i, \mathbf{R}_i \mid \mathbf{X}_i, \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \mathbf{R}_i, \boldsymbol{\theta}\right) \\ \times \Pr\left(\mathbf{R}_i \mid \mathbf{X}_i, \boldsymbol{\psi}\right).\end{aligned} \tag{3.12}$$

The PMM has desirable properties especially when NMAR (probability that a response will be missing depends on the $\mathbf{R}_i$ and $\mathbf{Y}_i$) situations are examined in the analysis (Wu and Bailey, 1988). For instance, where it is not substantively reasonable to consider non-responses as missing data, it may be desirable to limit the inferences to the subpopulation of patients whose responses are

observed. Thus, it is more meaningful to consider the distribution of Y_{ij} given $R_{ij} = 1$ ($R_{ij} = 1$ if subject is observed and 0 otherwise), rather than the marginal distribution of Y_{ij} (Rubin, 1976). Contrary to the selection model, $\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{X}_i, \mathbf{R}_i)$ is modeled directly from the pattern-mixture model.

One important feature of the pattern-mixture model (3.12) is that it fits different response model for each pattern of missingness such that the observed data is a mixture of patterns weighted by their respective probabilities of missing patterns. That is, the first component in the PMM (3.12), $\Pr(\mathbf{Y}_i \mid \mathbf{X}_i, \mathbf{R}_i, \boldsymbol{\theta})$ fits a response model for each pattern of missingness and $\Pr(\mathbf{R}_i \mid \mathbf{X}_i, \boldsymbol{\psi})$ represents drop out probability for each pattern. It follows that if there are U number of missingness patterns in a data set, then the marginal distribution of $\mathbf{Y}_i$ is a mixture of $\Pr(\mathbf{Y}_i \mid \mathbf{X}_i, \boldsymbol{\theta}) = \sum_{u=1}^{U} \Pr(\mathbf{Y}_i \mid \mathbf{R}_i = \mathbf{R}_i^u, \mathbf{X}_i, \boldsymbol{\theta}^u)\, \pi_u$, where $\pi_u = \Pr(\mathbf{R}_i = u \mid \mathbf{X}_i, \boldsymbol{\psi})$ and $\mathbf{R}_i$ counts the number of U patterns, $\boldsymbol{\theta}^u$ represents the parameters of marginal density $\Pr(\mathbf{Y}_i)$ in the u^{th} pattern. It can be observed that in the pattern-mixture model, parameters $\{\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^U\}$ can have different dimensions. A logistic model is often assumed for dropout probabilities and a linear mixed effected model (Laird and Ware, 1982) for the measurement process.

The pattern-mixture model (3.12) is well understood using the second MAR assumption. The second MAR assumption states that observations that would have been recorded for a patient in the future, given that the observed history of such a patient has the same statistical behavior irrespective of whether such patient dropout or does not dropout in the future. This feature of the pattern-mixture model makes it possible for multiple imputation to provide a practical approach to estimation and inferences. In addition, this feature provides a framework for the formulation of the pattern-mixture model with multiple imputation (Carpenter et al., 2013).

The pattern-mixture modeling framework was proposed by Wu and Bailey (1988); Wu and Carroll (1988); Wu and Bailey (1989) as approximate methods for making inferences about the time course of the continuous responses. Their approach was aimed at avoiding dependence on the complex model fitting in the selection modelling approach, earlier proposed by Wu and Bailey (1989). Their approach was based on method-of-moments type fitting of a linear model to the least squares slopes, given the dropout time and then averaging over the distribution of the dropout time. Apart from assuming that dropouts occur at discrete times, their approach does not provide any distributional assumptions on the event times. This approach was latter extended by Hogan and Laird (1997) to allow censored dropout times, which might occur when there is a group of patients who joined the trial at the latter period and temporary analyses can be carried out. Following later modification to the pattern-mixture model proposed by Wu and Bailey (1989), Follmann and Wu (1995*a*) generalized their conditional linear model to allow generalized linear models without any parametric assumption on the random effects. Some research related to pattern-mixture models can be found in Rubin (1977); Fitzmaurice and Laird (2000). Thijs et al. (2002) outlined a hierarchy for the

different ways of handling and fitting pattern-mixture models.

The pattern mixture models are rarely used for an arbitrary pattern of missingness because of the increase in the number of potential patterns. However, this is unlikely to occur when missingness is confined to dropout. The pattern-mixture model implies different distributions for each pattern of missingness or dropout. This implies that each deviation pattern will have its associated mean vector $\boldsymbol{\mu}_u$ and covariance matrix $\boldsymbol{\Sigma}_u$, where u represents deviation or dropout pattern. By construction, pattern-mixture models are under-identified or over-specified such that the data do not contain information on all the parameters specified by the model (Mallinckrodt et al. 2013; Molenberghs et al. 2014, pp.68).

The issue with estimating the inestimable parameters from the pattern-mixture model can be solved by using identifying restriction. Under such restriction, the parameters that cannot be estimated from the incomplete patterns are set equal to function of the parameters describing the distribution of other patterns. There are two main approaches to identifying pattern-mixture models. The first approach involves the use of outcome models that are sufficiently constrained so that such models can be identified within the different dropout patterns. However, the constraints that are required to make the model estimable in all the different dropout pattern implies the use of polynomial extrapolation which may be difficult to justify from substantive perspective (Molenberghs et al. 2014, pp.71).

The second approach, which is widely used to identifying pattern-mixture models relies on the idea of identifying restrictions. The basic idea behind identifying restriction is that one can either "borrow" the unidentifiable distributional information from the completers (Little, 1993), know as complete case missing values (CCMV) or from the nearest identified pattern (Mallinckrodt, Lin and Molenberghs, 2013), known as neighbouring case missing values (NCMV). An alternative identifying restriction approach is the available case missing value (ACMV); where the unavailable information is borrowed from the observed data.

A recent identifying restriction approach, which has been proven to be useful is to equate conditional distributions from the different treatment groups with the objective of representing patients in the study who deviate from the study protocol. However, whether such restriction will be intuitive is subject to the context of the study, particularly the objectives of the analysis, the nature of the outcome measurement and the action of the interventions (Albert et al., 2002; Carpenter et al., 2013). Analysis based on this identifying restriction has been developed by Little and Yau (1996) and then recently developed by Carpenter et al. (2013) for handling protocol deviations leading to missing data.

3.1.3 The shared-parameter modeling framework

In the shared-parameter modelling framework, the measurement and the missingness models share a common random effect or latent variable. The random effect captures the association between the measurements and missingness processes and also accounts for the correlation between the repeated measurement. This approach is often used to jointly modeled longitudinal measurements and time-to-event analysis of dropout.

The shared-parameter model (SPM) is defined as

$$\Pr\left(\mathbf{R}_i, \mathbf{Y}_i, \mathbf{b}_i \mid \mathbf{X}_i, \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \mathbf{b}_i, \boldsymbol{\theta}\right) \Pr\left(\mathbf{R}_i \mid \mathbf{X}_i, \mathbf{b}_i, \boldsymbol{\psi}\right) \Pr\left(\mathbf{b}_i \mid \mathbf{X}_i, \mathbf{G}\right), \tag{3.13}$$

where $\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \mathbf{b}_i, \boldsymbol{\theta}\right)$ is density function of the measurement process. The missingness process model is $\Pr\left(\mathbf{R}_i \mid \mathbf{X}_i, \mathbf{b}_i, \boldsymbol{\psi}\right)$ and $\Pr\left(\mathbf{b}_i \mid \mathbf{X}_i, \mathbf{G}\right)$ is the model for the random effects. $\mathbf{Y}_i$ and $\mathbf{R}_i$ are assumed to be conditionally independent, given the random effects $\mathbf{b}_i$. The equation (3.13) can alternatively be written as

$$\Pr\left(\mathbf{R}_i, \mathbf{Y}_i \mid \mathbf{X}_i, \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \int \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \mathbf{b}_i, \boldsymbol{\theta}\right) \Pr\left(\mathbf{R}_i \mid \mathbf{b}_i, \boldsymbol{\psi}\right) \Pr\left(\mathbf{b}_i \mid \mathbf{G}_i\right) \mathrm{d}\mathbf{b}_i. \tag{3.14}$$

Tsonaka et al. (2009) have demonstrated that shared-parameter model, by construction, implies a missing not at random (MNAR) mechanism. Thus we can write the conditional distribution

$$\Pr\left(\mathbf{R}_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m, \boldsymbol{\psi}\right) = \int \Pr\left(\mathbf{R}_i \mid \mathbf{b}_i, \boldsymbol{\psi}\right) \Pr\left(\mathbf{b}_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m\right) \mathrm{dY}_i^m, \tag{3.15}$$

where the random effects for the ith patient is assumed to be $\mathbf{b}_i \sim N\left(\mathbf{0}, \mathbf{G}\right)$. It can be observed that the response indicators depend on the missing response $\mathbf{Y}_i^m$ through the posterior distribution $\Pr\left(\mathbf{b}_i \mid \mathbf{Y}_i^o, \mathbf{Y}_i^m\right)$ of the random effects. This shows that the SPM is a joint model for non-ignorable missingness.

A shared-parameter model is defined by specifying a model $\Pr\left(\mathbf{Y}_i \mid \mathbf{b}_i\right)$ for the measurement process and a model $\Pr\left(\mathbf{R}_i \mid \mathbf{b}_i\right)$ for the missingness process. The most commonly used measurement model is the linear mixed effect model for the continuous longitudinal outcomes and a logistic regression model describing the probabilities of dropout.

Since the measurement and the dropout models (3.13) share common random effects, setting up a model that assumes that these models do not share common random effects (no linkage) is analogous to MAR mechanism and hence standard methods of analyses can be use to produce valid inferences. On the other hand, we may also assume that the measurement and the dropout models share a common random effect (linkage). This allows us to make alternative and plausible assumptions about the manner in which these random effects are shared between the measurement and the dropout models. This provides a formal framework for assessing sensitivity of the no linkage analysis results to the different assumptions about the linkage between the measurement and the

dropout models. This implies that the dependence of the measurement and the dropout models on the random effects forms the basis for sensitivity analyses.

Several authors have proposed the SPM for incomplete longitudinal data. The use of the shared-parameter model (Wu and Bailey, 1988; Wu and Carroll, 1988; Creemers et al., 2010) has been one approach to accounting for nonrandom missingness. In Wu and Carroll (1988) study of repeated measurements of lung function, they proposed the SPM. Their study assumed the linear mixed effect model (LMM) (Laird and Ware, 1982) with a random intercept and slope terms. The dropout process was modeled using a probit model for the censoring process. The LMM was linked with the censoring process model by including a patient's random slope as a covariate in the probit model for the censoring process. In this way, when the probit regression coefficient for the random slope is non-zero, it gives an indication that there is dependence between the measurement and the missingness processes.

The random effects in the SPM reflect patient's deviation from the mean estimates of the fixed effects. Each patient draws a random slope from a Gaussian distribution. This slope governs both the patient's expected rate of decline in the response and the probability of dropping out (see Albert and Follmann in Fitzmaurice et al., 2008, Ch. 19). They also pointed out that the shared-parameter models make assumptions which can only in part be justified, using the data. For instance, one can determine whether, for the observed data, change in response follows a linear function of time or a quadratic function of the log time, using standard diagnostics tools such as examining the residual correlation. On the contrary, changes in the response for the unobserved data, which are imposed by the SPM, cannot be verified.

The choice of measurement model depends on the type longitudinal data being analyzed. For normally distributed longitudinal data, the linear mixed model is often assumed. For discrete or dichotomous longitudinal data, the response model can be formulated as a generalized linear mixed effect model GLMM (Follmann and Wu, 1995*a*; Thomas et al., 1998; Albert and Follmann, 2000). The type of model for the missingness process depends on the type of missing data being considered. For instance, when the data are a discrete time to dropout then a geometric distribution can be used to model the missingness process (Mori et al., 1994). Mori et al. (1994) geometric model assumes that the number of observations per patient in the study is determined by his or her true rate of change. Several authors have proposed the SPM for the case where dropout is a continuous event time (Schluchter, 1992; De Gruttola and Tu, 1994; Tsiatis et al., 1995; Tsiatis and Davidian, 2001; Vonesh et al., 2002) and the logistic regression model is assumed for the probability of dropout.

3.1.4 Comparison of selection and pattern-mixture models

One major advantage of the selection models is that they directly modeled the marginal distribution of the complete data, which is often the usual focus of inferences in longitudinal study. Some statisticians prefer selection models since they offer straight forward formulation of the hypothesis about the non-response process (Diggle and Kenward, 1994; Diggle et al., 2002). Although assumptions about the non-responses are clear in selection models, what is less clear is how these translate into assumptions about the distributions of the unobserved outcomes. Also, because the model is over-specified (inadequate data to estimate some of the parameters in the model), identification comes from postulating unverifiable models for the dependence of the non-response process on the unobserved outcomes (Fitzmaurice et al., 2008, pp. 401). However, except in very simple cases, it can be very difficult to determine the identifying restrictions that must be placed on the model (Cole et al., 2005; Glonek, 1999). The selection models can be computationally demanding since estimation of parameters in the selection model model requires difficult integration over the distribution of the unobserved responses.

Pattern-mixture models are straightforward to fit as standard models that assume that non-response is ignorable. The pattern-mixture models are often over-specified or under-identified such that parameters describing the incomplete patterns cannot be estimated (Carpenter et al., 2007; Carpenter and Kenward, 2012; Mallinckrodt, Lin and Molenberghs, 2013). Identification is achieved by assuming unverifiable links among the distribution of the outcomes conditional on the missingness patterns. It is straightforward to determine identifying restriction in selection models than in pattern-mixture models as the number of patterns increased.

Both the selection and the pattern-mixture models specifications contain parameters that cannot be estimated without imposing arbitrary modeling assumptions about the distribution of such parameters or distributions describing the incomplete patterns or data (Carpenter et al., 2013).

3.2 Importance of conducting sensitivity analysis

As we have stated in this chapter, our sensitivity analyses are based on the joint models for non-ignorable missingness (NMAR). In this section, we explain why there is the need to consider various forms of sensitivity analyses methods when analyzing incomplete longitudinal data.

When analyzing incomplete longitudinal data, valid likelihood inference can be obtained when the data are missing at random (MAR) or when the data are missing completely at random (MCAR). However, when the data are missing not at random (non-ignorable missing data), valid inference can be obtained by modeling the missing data explicitly or implicitly. It is known that different

models for the data NMAR may produce identical fits of the data but with different implications for the unobserved data (Molenberghs et al., 2008; Sotto, 2009). This paradox may lead to different conclusions. Sensitivity analysis is therefore required to deal with this ambiguity in the modeling of data NMAR. There is therefore the need to assess the robustness of inferences under MAR assumption to alternative plausible assumptions under NMAR mechanism.

The US National Research Council (NRC, 2010) on prevention and treatment of missing data in clinical trials noted that "sensitivity analysis should be part of the primary reporting of findings from clinical trials ". Also, the Committee for Medicinal Products for Human Use (CHMP, 2010) guidelines on missing data in confirmatory clinical trials recommends that "sensitivity analyses should show how different assumptions influence the results obtained."

Recognizing the importance and, at the same time, difficulties associated with modeling incomplete data, especially in longitudinal clinical trials, the recent guideline and reports (CHMP, 2010; NRC, 2010) on treatment and prevention of missing data emphasize the importance of minimizing the amount of missing data. These guidelines also recommend carefully selecting primary analysis methods on the basis of assumptions regarding the missing data mechanism suitable for the study at hand. These recent guideline and reports also emphasize on the need to "stress-test" the results of the primary analysis under different sets of assumptions through a range of sensitivity analyses. The NRC guidelines recommend that sensitivity analyses should be planned to assess the impact of missing data on the study results. It should be noted that merely running additional analyses that make the same missing data assumptions is not useful. For instance, if a primary analysis assumes that the data are MAR, then a sensitivity analysis involving multiple imputation under the same assumption is uninformative. The number of sensitivity analyses conducted is not as important as the way in which the assumptions are varied (NRC, 2010; CHMP, 2010). In addition, "tipping point" (TP) analysis recommended by NRC is appealing for regulators. Tipping points are outcomes that result in a change of study conclusion. The TP analysis helps clinical reviewers in making judgment regarding treatment effect in the study (Smuk, 2015).

It is important to remind readers that after model fitting and sensitivity analysis, investigators or clinicians have to decide on how important the treatment effect is. However, there is no scientific agreement on how to convert information from a sensitivity analysis into a single decision about treatment effect. A viable approach is to draw inference under MAR mechanism and then investigate a set of sensitivity parameter values that would lead to overturning the conclusion from MAR. The results can be viewed as ambiguous if the inference about treatment effects could be overturned. Especially, values of the sensitivity parameter that are plausible (NRC, 2010).

Irrespective of the specific method taken to decision-making, the key issue is weighting the results, either formally or informally, from both the primary analysis and each alternative analysis by as-

sessing the plausibility of the assumptions made in conjunction with each analysis. The analyses should be given some degree of weight especially when the associated assumptions are viewed as being extreme. Also, the analyses should be given substantial weight when the associated assumptions are viewed as being comparably plausible to those for the primary analysis. This means that in situations in which there are alternative analyses as part of the sensitivity analysis that support contrary inferences to that of the primary analysis, if the associated assumptions are viewed as being fairly extreme, it would be reasonable to continue to support the inference from the primary analysis.

Recommendation 15 of the National Research Council (NRC) guidelines (NRC, 2010) on prevention and treatment of missing data recommend that sensitivity analyses should be part of the primary reporting of findings from clinical trials and examining sensitivity to the assumptions about the missing data mechanism should be a mandatory component of reporting. This guidelines noted that there are some often-used models for the analysis of missing data in clinical trials for which the form of a sensitivity analysis has not been fully developed in the literature. Also, the guidelines have provided principles for the broad development of sensitivity analyses. However, the guideline have not been prescriptive for many individual models. It is important that additional research be carried out so that methods of carrying out sensitivity analyses for all of the standard models are available.

In this thesis, our sensitivity analyses are based on the selection model, pattern-mixture model, and shared-parameter model frameworks (Rubin, 1976; Little, 1993; Diggle and Kenward, 1994; Robins et al., 1999; Molenberghs, Kenward and Goetghebeur, 2001; Verbeke et al., 2001; Diggle et al., 2002; Fitzmaurice et al., 2008). The sensitivity analyses considered include the local influence approach (Kenward, 1998; Verbeke et al., 2001; Shen et al., 2006; Thijs et al., 2000), and stress-testing methodology (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013), the pattern-mixture model with multiple imputation (Carpenter et al., 2013), and the sensitivity analysis for the generalized shared-parameter model. We explain the specification of each of these methods and state how each method allows us to perform sensitivity analysis.

We have stated in our previous discussion that there is no standard/universal approach for conducting sensitivity analysis to missing data. This is because the objective of clinical trial varies from trial to trial as well as the method of analysis. However, there are guidelines (NRC, 2010) on how appropriate sensitivity analysis may be conducted, given the trial setting/objective (that is the estimand of interest). Sensitivity analysis to missing data should consider various approaches that scientifically stress-test statistical inferences under a given primary analysis formulation by formulating missing data assumptions that deviate from that under the primary analysis assumption. The guidelines (CHMP, 2010; NRC, 2010) on missing data in confirmatory clinical trials recommend that sensitivity analyses should show how different assumptions influence the results

obtained. Just performing additional analyses that make the same missing data assumptions is not useful sensitivity analysis. The number of sensitivity analyses considered‘ is not as important as the way in which the assumptions are varied (NRC, 2010; CHMP, 2010).

Recognizing the importance and need for rigorous sensitivity analysis to missing data, we considered the three modelling frameworks (which make different assumptions about the missing data) for sensitivity analysis to the missing CD4 count data in the IMPI trial. Under each modelling framework, we point out the model specification that allows us to conduct sensitivity analysis in general, and in relation to the IMPI trial and other trial with similar settings. A careful study of these three main methods for conducting sensitivity analysis as well as their applications in trial settings offers opportunity (1) for proper understanding on which sensitivity analysis method to use and when, (2) to increase level of confidence that the reader of the report may have on the results, (3) to reader to decide on which of the methods he/she prefers to use to assess the relevance of the sensitivity analysis conducted, and (4) to the researcher to become an authority in modelling missing data.

Chapter 4

Sensitivity analyses for the selection model framework

4.1 Introduction

In this chapter, we assess sensitivity of statistical inferences to missing data assumptions. The sensitivity analyses approaches considered in this chapter are based on selection model framework (Diggle and Kenward, 1994) introduced in the Section 3.1.1. The sensitivity analyses considered are the local influence (Verbeke et al., 2001; Thijs et al., 2000, 2002; Shen et al., 2006) and the stress-testing (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013) approaches. The local influence approach enables us to assess the influence of potentially influential subjects on parameter estimates as well the dropout mechanism. The stress-testing approach, on the other hand, allows us to consider separate missingness models for each treatment arm, and then investigate the effect of different probabilities of dropout on parameter estimates and the dropout mechanism.

We describe the selection model formulation and explain the association between the measurement and dropout models, in the Diggle and Kenward (DK) selection model, that allows us to perform various forms of sensitivity analysis to nonrandom dropout. We point out the relevance of the DK selection model especially, the local influence and the stress-testing methodologies and explain how the missing data assumptions under these approaches are related to the assumptions about the unobserved CD4 count data. These approaches are then applied to the incomplete CD4 counts data and we discuss the results, comparing our findings in the local influence approach to the findings in the selection model approach and then findings in the stress-testing approach to the local influence approach. These comparisons offer more insight into the impact of various degrees of dropouts on the treatment effect. These analyses also allow us to assess plausibility of perturbation in the dropout mechanism and their implication on the treatment effect obtained. For instance, unrealistic perturbation in the dropout mechanism may overturn treatment effect or may not overturn the treatment effect. The unrealistic perturbations are expected to overturn the treatment effect and may have clinical explanation. However, explanation becomes difficult when unrealistic perturbations do not overturn the effect of treatment (NRC, 2010; Mallinckrodt, Roger, Chuang-Stein,

Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013).

Likelihood-based or Bayesian analyses can be used to produce valid inferences under MAR (Rubin, 1976; Molenberghs et al., 2008). However, one cannot rule out the possibility of data being NMAR Molenberghs et al. (2008). Various authors have proposed several statistical methods for the NMAR mechanism (Diggle and Kenward, 1994) and the sensitivity of such methods has been studied by several authors (Molenberghs et al., 2008; Diggle and Kenward, 1994). Several authors have explained such sensitivities and subsequently proposed informal and formal methods for sensitivity analysis (Kenward, 1998; Verbeke et al., 2001). Molenberghs et al. (2008) discussed the need to perform sensitivity analysis to compare the sensitivity of inferences under primary MAR analyses to alternative plausible assumptions under NMAR models for sensitivity analyses (Thijs et al., 2000; Molenberghs et al., 2008; Diggle and Kenward, 1994; Verbeke et al., 2001; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013).

Several authors proposed the NMAR selection models for analyzing incomplete longitudinal data (Molenberghs et al., 1997; Van Steen et al., 2001; Jansen et al., 2003; Rubin, 1976; Little, 1995). The DK selection modeling framework factorizes the joint distribution of the measurement and the missingness mechanisms into marginal distribution of the measurement process and the conditional distribution of the missingness process, given the response. The Diggle and Kenward (1994) selection model was applied to the milk protein data (Verbyla and Cullis, 1990) where they concluded that the dropout mechanism was nonrandom.

The DK selection model received several criticisms ranging from computational instability (Shen et al., 2006) and the use of such models (Thijs et al., 2000). In particular, conclusions based on such models have been questioned in terms of their reliability because of their strong but untestable assumptions about the NMAR mechanism. The nature of the incompleteness was questioned since it was due to design reason but not a genuine dropout (Diggle and Kenward, 1994; Kenward, 1998; Thijs et al., 2000). Several authors have argued that such models should not be used to conclusively determine whether or not the dropout mechanism is nonrandom. Little (1995) argued that parameter estimates depend on normal assumptions and correct specification of the model, whereas Laird's discussion to Diggle and Kenward (1994) noted that estimating the untestable assumptions can only be achieved by making modeling assumptions about the dropout mechanism. Thus the consequences of model misspecification will probably be far more severe when the dropout mechanism is non-ignorable.

Also, Rubin in his discussion to Diggle and Kenward (1994) selection model argued that inferences for the data parameters generally depend on the assumed missingness mechanism; which implies greater sensitivity of inference to a reasonable model specification. Molenberghs et al. (1997) also argued that conclusions are conditional on the appropriateness of the assumed model, which is

untestable. As a result, the view has emerged that such analyses should be placed within sensitivity analysis framework.

Verbeke and Molenberghs (2009) sensitivity analysis was based on the DK selection modeling framework. They showed that excluding a small number of measurement errors considerably changes the likelihood ratio test statistics for the MAR null hypothesis. Kenward (1998) re-analyzed the milk protein data dataset using the DK selection model. Kenward demonstrated that removing two unusual observations from the analyses changed the nonrandom dropout mechanism towards random dropout mechanism. This implies that the presence of these two subjects in Diggle and Kenward (1994) analyses lead to the conclusion that the dropout mechanism was nonrandom. Molenberghs, Verbeke, Thijs, Lesaffre and Kenward (2001) confirmed Kenward (1998) results using influence diagnostics.

These considerations motivate the use of influence diagnostics to conduct sensitivity analysis to the incomplete CD4 counts measurements from the IMPI clinical trial (Mayosi et al., 2014). As noted earlier, the unobserved CD4 counts measurements were due to administrative reasons. Hence, it is unlikely that the probability of dropout will depend on the unobserved CD4 counts measurements. The influence analysis considered in this study focuses on the local influence approach to identify patients with unusual observations that may lead missingness mechanism towards NMAR (Jansen et al., 2006). The local influence considered here is based on perturbing the non-random part of the DK selection model to enable us to assess the influence of such perturbation on the MAR mechanism. A careful study of such patients may lead to an appropriate level of confidence being placed on the proposed MAR primary analysis.

4.2 Methodology

This section describes the Diggle and Kenward (1994) methodology for handling nonrandom dropout. We have already introduced this methodology in Section 3.1.1. In order to estimate parameters $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ we require a model for the measurement $\Pr(\mathbf{Y}_i \mid \boldsymbol{\theta})$ and the dropout $\Pr(D_i \mid \mathbf{Y}_i, \boldsymbol{\psi})$ processes respectively.

4.2.1 Model for the measurement process

The linear mixed effects model (2.3) (Laird and Ware, 1982), described in the Section 2.5.1, is assumed for the measurement process.

4.2.2 Model for the dropout process

The model for the dropout process is based on a logistic regression for the probability of dropout at the jth occasion, given that the subject was observed at the previous occasion and depends on the observed history $\mathbf{h}_{ij}$, and the current observation Y_{ij} but not future outcomes (Diggle and Kenward, 1994; Verbeke et al., 2001). Let $g\left(\mathbf{h}_{ij}, Y_{ij}\right)$ denote the dropout probability. The probability of dropout $\Pr\left(D_i \mid \mathbf{Y}_i, \boldsymbol{\psi}\right)$ under the SeM (3.7) assumes that $g\left(\mathbf{h}_{ij}, Y_{ij}\right)$ satisfies (Diggle and Kenward, 1994)

$$\operatorname{logit}\left[g\left(\mathbf{h}_{ij}, Y_{ij}, \boldsymbol{\psi}\right)\right] = \operatorname{logit}\left[\Pr\left(D_i = j \mid D_i \geq j, \mathbf{Y}_i, \boldsymbol{\psi}\right)\right] = \mathbf{h}_{ij}'\boldsymbol{\psi} + \omega Y_{ij}, \tag{4.1}$$

where $\boldsymbol{\psi}' = (\psi_0, \psi_1, \omega)$, ψ_0 is the intercept, ψ_1 and ω are regression coefficients describing the previously observed outcomes $\mathbf{h}_{ij}' = (Y_{i1}, Y_{i2} \ldots, Y_{ij-1})$ and the current outcome Y_{ij} respectively (Diggle and Kenward, 1994; Verbeke et al., 2001). It can be noted that the response variable Y_{ij} $(\mathbf{Y}_i)$ in the linear mixed effect model (2.3) is a covariate in the dropout model (4.1). The measurement and the dropout models are linked by the parameter estimate ω. This feature of the dropout model (4.1) is more convenient for some sensitivity analyses (Verbeke et al., 2001; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013). For instance, if $\omega = 0$, the missing data mechanism is assumed to be MAR and if $\omega \neq 0$, the missing data mechanism is NMAR and thus enables us to investigate the sensitivity of inferences under MAR mechanism compared to various plausible assumptions under NMAR models. The model (4.1) provides the building blocks for the dropout process $\Pr(d_i \mid \psi_0, \psi_1, \omega)$.

Suppressing the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$, the dropout model can be written as (Diggle and Kenward, 1994; Verbeke et al., 2001)

$$\Pr\left(d_i \mid \mathbf{Y}_i\right) = \begin{cases} \prod_{j=2}^{n_i}\left[1 - g\left(\mathbf{h}_{ij}, Y_{ij}\right)\right] & \text{for a completer} \quad (d_i = n_i + 1) \\ \prod_{j=2}^{d_i - 1}\left[1 - g\left(\mathbf{h}_{ij}, Y_{ij}\right)\right] g\left(\mathbf{h}_{id}, Y_{id}\right) & \text{for a dropout} \quad (d_i = d < n_i), \end{cases} \tag{4.2}$$

where the g-factors follow from the model (4.1).

It can be noted that the response variable Y_{ij} in the linear mixed effect model (2.3) is a covariate in the dropout model (4.1). That is, the measurement and the dropout models are linked by the parameter estimate ω. This feature of the dropout model (4.1) is more convenient for conducting sensitivity analyses (Verbeke et al., 2001; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Molenberghs et al., 2014; Shen et al., 2006; Thijs et al., 2000, 2002; Kenward, 1998). For instance, if $\omega = 0$, the missing data mechanism is assumed to be MAR and hence dropout no longer depends on the Y_{ij}, and all parameters can be estimated using standard software packages. However, if $\omega \neq 0$, the missing data mechanism is NMAR and thus ables us

to investigate the sensitivity of statistical inferences under MAR mechanism to various plausible assumptions under NMAR models.

In practice, some subjects may dropout randomly while others dropout nonrandomly (Thijs et al., 2000). However, model (4.1) cannot accommodate this possibility as ω does not vary across subjects. Verbeke and colleagues Verbeke et al. (2001) noted that a dropout model may be nonrandom only because one or few influential subjects have affected the analysis and subsequently proposed the local influence approach to detect subjects who have unusual observations that are likely to distort the key conclusions from the model (Verbeke et al., 2001).

4.3 Sensitivity analysis approaches

In this section, we discuss the two sensitivity analysis approaches based on the DK SeM (Diggle and Kenward, 1994). We discuss the local influence analysis in Section 4.3.1 and the stress-testing analysis in Section 4.3.2.

4.3.1 Influence analysis as sensitivity tool

This section describes the influence analysis approaches with main focus on the local influence sensitivity analysis approach to nonrandom dropout (Verbeke et al., 2001). The other influence approach is the global (Cook and Weisberg, 1982; Chatterjee and Hadi, 2009). Whereas these approaches have similar objectives of identifying patients who have unusual observations and thus are likely to distort key model's conclusions, they differ in terms of approach to assessing influence. These approaches differ in terms of perturbation schemes used for assessing the influence of, potentially, influential patients on the estimation of the model parameters.

The most commonly used perturbation scheme is the global influence case-deletion, in which the impact of an influential patient is studied by completely removing such patients from the analysis. This sensitivity-analysis tool starts from case deletion and is based on the difference in log-likelihood between the model fitted to the entire dataset (Molenberghs et al., 2014, Sec.16.5.2). Although the global influence analysis is conceptually simple, computationally straightforward and well studied, it cannot be extended to a more general settings. This is because the global influence requires N-dimensional model fits (Thijs et al., 2000). There is, however, one fundamental problem. That is, under the global influence, it is not clear as to how to assess the influence that can be attributed to a specific cause, since by removing a patient, all kinds of influences resulting from such a patient are lumped together. That is, the influence is likely to be caused by different mechanisms and such an investigation cannot be done with the case-deletion method because it is not possible to disentangle

the various sources of influence (Fitzmaurice et al., 2008, Ch. 22, Sec. 22.7). These limitations can be avoided by using the local influence approach (Verbeke et al., 2001). The local influence approach (Verbeke et al., 2001) allows one to assess direct and indirect influences on the dropout and measurement model parameters, stemming from perturbing the random dropout model in the direction of nonrandom dropout. This can be achieved by perturbing the nonrandom dropout part of the model for the dropout.

The local influence approach involves perturbing some components of the model and then assessing the impact of such perturbations on the model parameters of interest. Examples of such perturbations include (1) perturbing the variances, or the fixed effects components of the measurement model (2.3), and (2) making different assumptions about the missingness mechanism in the dropout model (4.1) (Verbeke et al., 2001; Verbeke and Molenberghs, 2009), or (4) changing the distributional assumption of the model (Kenward, 1998). In this study, our perturbations scheme is based on the dropout model specified by the DK selection model. Perturbation was achieved by perturbing the NMAR part (ω) of the model (4.1) around the null model (MAR).

One important property of the local influence approach is that one is able to attribute the influence of influential patients to their specific characteristics by inspecting the index plots of the components of the model (Lesaffre and Verbeke, 1998; Verbeke et al., 2001; Shen et al., 2006). In this study, the components inspected include (1) the fixed effects, (2) variance of the random effects, (3) parameters describing the dropout mechanism (4) the maximum influence measure and (5) total influence. The total influence is the overall influence from components 1, 2 and 3.

4.3.1.1 Local influence analysis

This section describes the local influence methodology (Verbeke et al., 2001; Shen et al., 2006). The local influence approach involves perturbing some components of the model and then assessing the impact of such perturbations on the model parameters of interest. Several authors (Verbeke et al., 2001; Thijs et al., 2000) have investigated the sensitivity analysis of dropout model parameter estimates, with respect to the dropout model assumption. Verbeke et al. (2001) considered the perturbed version of the dropout model (4.1) which is defined as

$$\text{logit}\left[g\left(\mathbf{h}_{ij}, Y_{ij}, \boldsymbol{\psi}\right)\right] = \text{logit}\left[\Pr\left(D_i = j \mid D_i \geq j, \mathbf{Y}_i, \boldsymbol{\psi}\right)\right] = \mathbf{h}_{ij}'\boldsymbol{\psi} + \omega_i Y_{ij}, \tag{4.3}$$

where ω is now replaced by ω_i. The term ω_i is not a parameter which needs to be estimated but is a local and subject-specific perturbations around the null model. The null model in this study will be MAR, which corresponds to setting $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)' = \mathbf{0}$ in equation (4.3), and thus enables studying the effect of how small perturbations, in the direction of NMAR ($\boldsymbol{\omega} \neq \mathbf{0}$), can have a large impact on key model components, such as treatment effect, variance components, or dropout

parameters. This can be achieved by constructing influence measures (Cook, 1986; Verbeke et al., 2001).

4.3.1.2 Local influence measures

Following George Box's famous remark that all statistical models are wrong, but some are useful, Cook (1986) used this idea to motivate his assessment of the local influence approach. Cook (1986) commented that more confidence can be put in a model which is relatively stable under small modifications.

The idea behind the local influence approach is to investigate how the results of an analyses change under several perturbations of the model. Cook (1986) proposed to study the influence of potentially influential subjects using the likelihood displacement. Beckman et al. (1987) used the same approach to detect influential observation in the linear mixed effect model (LMM) (Laird and Ware, 1982). Beckman et al. (1987) paper assessed the effect of perturbing the error variances, the random effect variances and the response vector. Christensen et al (1992) proposed the global influence approach for the same LMM as in Beckman et al. (1987). However, neither of their approaches was developed for repeated measures. This means that emphasis was placed on the influence of single observations.

Waternaux et al. (1989) proposed several procedures to detect outliers for the linear mixed effect model (LMM). However, their proposals are based on the global influence. Lesaffre and Verbeke (1998) introduced the local influence approach in the context of the LMM for longitudinal data. They showed that the local influence can be used to detect influential subjects in a longitudinal data analysis. Lesaffre and Verbeke (1998) noted that the method of local influence is particularly useful for the LMM because the ML and REML estimations produced for LMM are very sensitive to outlying observation and thus have the potential to detect such observations. One advantage of their approach is that the local influence diagnostics can be expressed analytically and can often be decomposed into interpretable components. These components allow us to determine whether a particular patient/observation is influential in the random or the fixed components of the LMM model. In other words, it offers insights into the reason why some patients are more influential than others. However, Lesaffre and Verbeke (1998) approach is limited to longitudinal data without missing data. This leads to the need for local influence approaches that could handle incomplete longitudinal data.

Verbeke et al. (2001) proposed a local influence approach to incomplete longitudinal data. They also deduced analytic expressions for the resulting local influence diagnostics and showed how these diagnostics can often be decomposed into interpretable components. Verbeke et al. (2001) proposed to study the effect that the non-randomness of dropout has on the model parameters of interest which are often the fixed effects and the variance components. This can intuitively be achieved by

considering the dropout model (4.3).

We now review the key concepts of the local influence (Cook, 1986). Many of the concepts used in this approach are taken from differential geometry (O'neill, 2006).

Let $\boldsymbol{\gamma}' = (\boldsymbol{\theta}', \boldsymbol{\psi}')'$ denote an s-dimensional vector grouping the parameters of the measurement model and the dropout model. Let $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\rho}, \boldsymbol{\sigma})$ and $\boldsymbol{\ell}(\boldsymbol{\gamma} \mid \boldsymbol{\omega}) = \sum_{i=1}^{N} \ell_i(\boldsymbol{\gamma} \mid \boldsymbol{\omega})$ denote the log-likelihood function corresponding to the dropout model (4.3), in which $\ell_i(\boldsymbol{\gamma} \mid \boldsymbol{\omega})$ denotes the log-likelihood function of the i^{th} subject's contribution to $\boldsymbol{\ell}(\boldsymbol{\gamma} \mid \boldsymbol{\omega})$. For $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, \ldots, 0)'$, $\boldsymbol{\ell}(\boldsymbol{\gamma} \mid \boldsymbol{\omega}_0)$ is the log-likelihood function corresponding to MAR dropout model (O'neill, 2006). Now let $\hat{\boldsymbol{\gamma}}$ denoting the maximum likelihood estimator of $\boldsymbol{\gamma}$, obtained by maximizing $\boldsymbol{\ell}(\boldsymbol{\gamma} \mid \boldsymbol{\omega}_0)$ and $\hat{\boldsymbol{\gamma}}_\omega$ denote the maximum likelihood estimator of $\boldsymbol{\gamma}$, obtained by maximizing $\boldsymbol{\ell}(\boldsymbol{\gamma} \mid \boldsymbol{\omega})$. Similar estimates give an indication that the parameter estimates are robust with respect to perturbations of the MAR dropout model in the direction of NMAR dropout model, whereas a large difference between the estimates indicates that the estimation procedure is very sensitive to such perturbations.

Cook (1986) proposed measuring the distance between $\hat{\boldsymbol{\gamma}}_\omega$ and $\hat{\boldsymbol{\gamma}}$ by using likelihood displacement $\text{LD}(\boldsymbol{\omega})$, defined by $\text{LD}(\boldsymbol{\omega}) = 2\left[\boldsymbol{\ell}(\hat{\boldsymbol{\gamma}} \mid \boldsymbol{\omega}_0) - \boldsymbol{\ell}(\hat{\boldsymbol{\gamma}}_\omega \mid \boldsymbol{\omega})\right]$, which takes into account the variability in $\hat{\boldsymbol{\gamma}}$. It is useful to view such a graph as a geometric surface formed by the $(N+1)$-dimensional vector $\boldsymbol{\xi}(\boldsymbol{\omega}) = \big(\boldsymbol{\omega}', \text{LD}(\boldsymbol{\omega})\big)'$ as $\boldsymbol{\omega}$ varies through $\boldsymbol{\Omega}$. Cook (1986) proposed the local influence measure that uses normal curvatures $C_{\tilde{\mathbf{a}}}$ of $\boldsymbol{\xi}(\boldsymbol{\omega})$ in $\boldsymbol{\omega}_0$, in the direction of some N-dimensional vector $\tilde{\mathbf{a}}$, of unit length. Cook (1986) showed that $C_{\tilde{\boldsymbol{a}}} = 2 \mid \tilde{\boldsymbol{a}}'\boldsymbol{M}'\boldsymbol{\Phi}^{-1}\boldsymbol{M}\tilde{\boldsymbol{a}} \mid$, where $\boldsymbol{M}_i$ is the s-dimensional vector defined as $\boldsymbol{M}_i = \frac{\partial^2 \ell_i(\boldsymbol{\gamma}|\omega_i)}{\partial \omega_i \partial \boldsymbol{\gamma}} \Big|_{\boldsymbol{\gamma}=\hat{\boldsymbol{\gamma}}, \omega_i = 0}$.

It follows that $\boldsymbol{M}$ is the $s \times N$ matrix with $\boldsymbol{M}_i$ as its ith column. $\boldsymbol{\Phi}$ denotes the second derivative of $\boldsymbol{\ell}(\boldsymbol{\gamma} \mid \boldsymbol{\omega}_0)$ evaluated at $\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}$, where $\hat{\boldsymbol{\gamma}}$ is maximum likelihood estimate of $\boldsymbol{\gamma}$. There are several ways in which $C_{\tilde{\boldsymbol{a}}}$ can be used to study $\boldsymbol{\xi}(\boldsymbol{\omega})$, each corresponding to a specific choice of unit vector $\tilde{\boldsymbol{a}}$ Verbeke et al. (2001). One choice of $\tilde{\boldsymbol{a}}$ is a vector $\tilde{\boldsymbol{a}}_i$ containing 1 in the ith position corresponding to the perturbation of the ith subject only and zero elsewhere (Thijs et al., 2000; Verbeke et al., 2001). In this way one is able to study the influence of allowing the ith subject to dropout nonrandomly while the others dropped out randomly. It follows from $C_{\tilde{\boldsymbol{a}}}$ that the corresponding local influence measure C_i for each subject is defined by $C_i = 2 \mid \boldsymbol{M}_i'\Phi^{-1}\boldsymbol{M}_i \mid$. The $\tilde{\boldsymbol{a}}_{\max}$ of the maximal curvature $C_{\max}$ shows how to perturb the null model (MAR) to obtain the largest local influence, which is the largest eigenvalue of $-2\boldsymbol{M}_i'\Phi^{-1}\boldsymbol{M}_i$ and $\tilde{\boldsymbol{a}}_{\max}$ is the corresponding eigenvector.

4.3.1.3 Local influence application under SeM framework

This section describes application of the local influence approach to the DK selection model (Diggle and Kenward, 1994). Let $\ell_{i\boldsymbol{\omega}}$ denote the log-likelihood function of the perturbed model (4.3). From

the dropout model (4.2), Verbeke et al. (2001) showed that the log likelihood function for completers (subjects who completed the study) is defined as

$$\ell_{i\boldsymbol{\omega}} = \ln \Pr\left(\mathbf{Y}_i\right) + \sum_{j=2}^{n_i} \ln\left[1 - g\left(\mathbf{h}_{ij}, Y_{ij}\right)\right],$$

whereas the log-likelihood function for the dropouts is defined as

$$\ell_{i\boldsymbol{\omega}} = \ln \Pr\left(Y_{i1}, \ldots, Y_{iD_i-1}\right) + \sum_{j=2}^{D_i-1} \ln\left[1 - g\left(\mathbf{h}_{ij}, Y_{ij}\right)\right] + \ln \int \Pr\left(Y_{iD_i} \mid Y_{i1}, \ldots, Y_{iD_i-1}\right) g\left(\mathbf{h}_{iD_i}, Y_{iD_i}\right) dY_{iD_i}.$$

For the completers (no dropout), Verbeke et al. (2001); Verbeke and Molenberghs (2009) showed that

$$\left.\frac{\partial^2 \ell_{i\boldsymbol{\omega}}}{\partial \boldsymbol{\theta} \partial \omega_i}\right|_{\omega_i=0} = \mathbf{0}$$

and

$$\left.\frac{\partial^2 \ell_{i\boldsymbol{\omega}}}{\partial \boldsymbol{\psi} \partial \omega_i}\right|_{\omega_i=0} = -\sum_{j=2}^{n_i} \mathbf{h}_{ij} Y_{ij} g\left(\mathbf{h}_{ij}\right)\left[1 - \lambda\left(\mathbf{h}_{ij}\right)\right].$$

and for the incomplete cases (dropouts)

$$, \left.\frac{\partial^2 \ell_{i\boldsymbol{\omega}}}{\partial \boldsymbol{\theta} \partial \omega_i}\right|_{\omega_i=0} = \left[1 - g\left(\mathbf{h}_{\mathbf{iD_i}}\right)\right] \frac{\partial \lambda\left(Y_{iD_i} \mid \mathbf{h}_{iD_i}\right)}{\partial \boldsymbol{\theta}}$$

and

$$\left.\frac{\partial^2 \ell_{i\boldsymbol{\omega}}}{\partial \boldsymbol{\psi} \partial \omega_i}\right|_{\omega_i=0} = -\sum_{j=2}^{D_i-1} \mathbf{h}_{ij} Y_{ij} g\left(\mathbf{h}_{ij}\right)\left[1 - g\left(\mathbf{h}_{ij}\right)\right] - \mathbf{h}_{id}\left(Y_{iD_i} \mid \mathbf{h}_{iD_i}\right) g\left(\mathbf{h}_{iD_i}\right)\left[1 - g\left(\mathbf{h}_{iD_i}\right)\right]. \tag{4.4}$$

The above expressions are evaluated at $\hat{\boldsymbol{\gamma}}$, and $g\left(\mathbf{h}_{ij}\right) = g\left(\mathbf{h}_{ij}, Y_{ij}\right) \mid \omega_i = 0$ corresponds to the MAR version of the dropout model. Computational details for the above expressions can be found in Verbeke and Molenberghs (2009). In expression (4.4), we make use of the conditional mean $\lambda\left(Y_{iD_i} \mid \mathbf{h}_{iD_i}\right)$. The functions $g(.)$ and $\lambda(.)$ are related in a sense that the response variable in $\lambda(.)$ is a covariate in $g(.)$. Let $V_{i,11}$ denote the predicted variance matrix for the observed vector $\left(Y_{i1}, \ldots, Y_{iD_i-1}\right)'$, $V_{i,22}$ be the predicted variance for the missing observation Y_{iD_i}, and $V_{i,12}$ the vector of the predicted covariances between the elements of the observed vector and the missing observations. Then it follows from the linear mixed effects model (2.3) that the conditional expectation in expression (4.4), for the observation at dropout, given the history of the subject is given

by (Verbeke and Molenberghs, 2009; Verbeke et al., 2001)

$$\lambda\left(Y_{iD_i} \mid \mathbf{h}_{iD_i}\right) = \lambda\left(Y_{iD_i}\right) + V_{i,21}V_{i,11}^{-1}\left[\mathbf{h}_{iD_i} - \lambda\left(\mathbf{h}_{iD_i}\right)\right]. \tag{4.5}$$

The derivatives of (4.5) with respect to the measurement model parameters are

$$\frac{\partial\lambda\left(Y_{iD_i} \mid \mathbf{h}_{iD_i}\right)}{\partial\boldsymbol{\beta}} = \mathbf{x}_{iD_i} - V_{i,21}V_{i,11}^{-1}\mathbf{X}_{iD_i-1}$$

and

$$\begin{aligned}\frac{\partial\lambda\left(Y_{iD_i} \mid \mathbf{h}_{iD_i}\right)}{\partial\boldsymbol{\alpha}} &= \left[V_{i,11}^{-1}\frac{\partial V_{i,21}}{\partial\boldsymbol{\alpha}} - V_{i,21}V_{i,11}^{-2}\frac{\partial V_{i,11}}{\partial\boldsymbol{\alpha}}\right]\left[\mathbf{h}_{iD_i} - \lambda\left(\mathbf{h}_{iD_i}\right)\right] \\ &= \left[\frac{\partial V_{i,21}}{\partial\boldsymbol{\alpha}} - V_{i,21}V_{i,11}^{-1}\frac{\partial V_{i,11}}{\partial\boldsymbol{\alpha}}\right]V_{i,11}^{-1}\left[\mathbf{h}_{iD_i} - \lambda\left(\mathbf{h}_{iD_i}\right)\right]\end{aligned}$$

where $\mathbf{x}'_{iD_i}$ is the D_i^{th} row of $\mathbf{X}_i$, and $\mathbf{X}_{i,D_i-1}$ indicates the first $D_i - 1$ rows of $\mathbf{X}_i$ and $\boldsymbol{\alpha}' = (\boldsymbol{\sigma}', \boldsymbol{\rho}')'$ is a parameter vector for the variance components. The parameter vector $\boldsymbol{\theta}$ in the measurement model is often of primary interest. Because $\boldsymbol{\Phi}$ is block-diagonal with blocks $\boldsymbol{\Phi}(\boldsymbol{\theta})$ and $\boldsymbol{\Phi}(\boldsymbol{\psi})$, the total influence is given by $C_{\tilde{\boldsymbol{a}}}(\boldsymbol{\gamma}) = C_{\tilde{\boldsymbol{a}}}(\boldsymbol{\theta}) + C_{\tilde{\boldsymbol{a}}}(\boldsymbol{\psi})$ for any unit vector $\tilde{\boldsymbol{a}}$ where

$$C_{\tilde{a}}(\boldsymbol{\theta}) = -2\tilde{\boldsymbol{a}}'\left[\frac{\partial^2\ell_{i\boldsymbol{\omega}}}{\partial\boldsymbol{\theta}\partial\omega_i}\bigg|_{\omega_i=0}\right]'\boldsymbol{\Phi}^{-1}(\boldsymbol{\theta})\left[\frac{\partial^2\ell_{i\boldsymbol{\omega}}}{\partial\boldsymbol{\theta}\partial\omega_i}\bigg|_{\omega_i=0}\right]\tilde{\boldsymbol{a}}$$

and

$$C_{\tilde{\boldsymbol{a}}}(\boldsymbol{\psi}) = -2\tilde{\boldsymbol{a}}'\left[\frac{\partial^2\ell_{i\boldsymbol{\omega}}}{\partial\boldsymbol{\psi}\partial\omega_i}\bigg|_{\omega_i=0}\right]'\boldsymbol{\Phi}^{-1}(\boldsymbol{\psi})\left[\frac{\partial^2\ell_{i\boldsymbol{\omega}}}{\partial\boldsymbol{\psi}\partial\omega_i}\bigg|_{\omega_i=0}\right]\tilde{\boldsymbol{a}}$$

are evaluated at $\boldsymbol{\gamma} = \hat{\boldsymbol{\gamma}}$.

Illustration of the local influence approach in a special case of three measurement occasions, using the three-dimensional version of the linear mixed effects model (2.3), can be found in Verbeke et al. (2001) and Verbeke and Molenberghs (2009). We compute the following normal curvatures in the direction of the unit vector $\tilde{\boldsymbol{a}}_i$ containing 1 in the ith position and zero elsewhere: $C_i(\boldsymbol{\gamma})$, $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\beta})$, $C_i(\boldsymbol{\alpha})$ and $C_i(\boldsymbol{\psi})$, as well as the normal curvature in the direction of $\tilde{\boldsymbol{a}}_{\max}$ of the normal curvature $C_{\max}$.

4.3.2 Stress-testing application to the SeM

In this section, we describe the stress-testing methodology for sensitivity analysis to nonrandom drop out. To assess the impact of different probabilities of dropout (at the respective treatment arms) on parameter estimates and the dropout mechanism, we consider separate missingness models for each treatment arm (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013). The probability of dropout was modelled using a logistic regression model specified under the DK SeM (Diggle and Kenward, 1994). We choose the DK logistic regression model because the current outcome is a covariate in this regression model and hence allows us to assess

the odds of dropout at a scheduled visit. We achieve this using the DK logistic regression model that fits the log odds of dropout as a function of scheduled visit time, separate intercepts $\boldsymbol{\psi}_0 = (\psi_{0p}, \psi_{0a})$ for each treatment arm, separate parameter estimates $\boldsymbol{\psi}_1 = (\psi_{1p}, \psi_{1a})$, and $\omega = (\omega_p, \omega_a)$ for the previous and the current outcomes respectively, where p and a refer to placebo and prednisolone arms respectively.

The parameter ω is of particular interest since it is the NMAR part of the model and hence sensitivity analysis should be based on a plausible range of values for ω_{p} and ω_{a}. These values assess the increase in the log odds of the dropout probability per unit increase in the outcome. Fitting a separate missingness model for each group allows for different departures from MAR for the prednisolone and placebo arms and thus allows us to determine the dropout rate in each arm which is likely to yield favorable or unfavorable treatment effects.

4.4 Analyses of the CD4 data

In this section, we specified the model and then described the effect and covariance parameters in the model. We also applied the sensitivity analysis approaches to the CD4 count data and then discussed the results obtained. In this section, the CD4 count data presented in Section 2.6 were analyzed under MCAR, MAR and NMAR. In the measurement model (2.3), we included an intercept, and assumed as fixed effects the following covariates: prednisolone (which takes the value of 1 for subjects randomized to prednisolone and 0 if the subject was randomized to placebo), time (months), age, whether on ART or not at each scheduled visit (1 if the subject received ART and 0 if subject did not receive ART), and interactions between prednisolone and time, and prednisolone and ART. Age and time were continuous variables. The ART is a time-varying covariate.

The trial was a cardiology trial but most of the patients were HIV positive, so HIV-related data were collected and we assessed the effects of ART on CD4 count changes over time and ART interaction with prednisolone. Further, our initial analyses showed that the confidence intervals for the parameter estimates of these variables do not vary significantly with overlapped confidence intervals except the intercept which varies significantly with non-overlapped confidence intervals. Also, the variance of the random effect for the intercept is very large while that of the other parameter estimates are almost zero. This suggest a random intercept model for the CD4 count data. Hence, our fitted linear mixed effects model (2.3) is defined as

$$\begin{aligned}\sqrt{\mathrm{CD4}_{ij}} = \beta_0 &+ \beta_1 \times \mathrm{prednisolone}_i + \beta_2 \times \mathrm{month}_j \\ &+ \beta_3 \times \mathrm{prednisolone}_i \times \mathrm{month}_{ij} + \beta_4 \times \mathrm{ART}_{ij} \\ &+ \beta_5 \times \mathrm{prednisolone}_i \times \mathrm{ART}_{ij} + \beta_6 \times \mathrm{Age}_i \\ &+ \mathrm{b}_i + \epsilon_{ij},\end{aligned} \tag{4.6}$$

where $\sqrt{\text{CD4}_{ij}}$ is the square root of CD4 counts for ith patient at the jth visit, for $i = 1, \ldots, N$ and $j = 1, \ldots, n_i$, b_i represents the patient-specific random effect, and ϵ_{ij} is the residual error. It is assumed that b_i and ϵ_{ij} are independently distributed as $\text{b}_i \sim N(0, \sigma_b^2)$ and $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$ respectively. We fitted the models using SAS software (version 9.5). The SAS programs can be found in **Appendix** A.

4.4.1 Results from the SeM

Table 4.1 shows that the statistical inferences under MAR assumptions are robust to the NMAR assumptions. There is no significant prednisolone effect and prednisolone and ART do not interact. CD4 count levels for subjects increased with increasing time and those on ART have higher CD4 count levels compared to those who are not on ART. There is no significant prednisolone-time interaction and older subjects have lower CD4 count levels than younger subjects.

Table 4.1: Parameter estimates for different missingness models fitted to the CD4 count data.

	MCAR			MAR			NMAR		
Parameter	Est.	s.e	p-value	Est.	s.e	p-value	Est.	s.e	p-value
Measurement model									
Time	0.27	0.10	0.008	0.27	0.10	0.008	0.25	0.12	0.042
Prednisolone	0.33	0.89	0.711	0.33	0.90	0.711	0.29	0.91	0.747
Prednisolone × Time	-0.08	0.15	0.598	-0.08	0.15	0.598	-0.08	0.15	0.579
ART	3.10	0.57	<0.000	3.10	0.57	<0.000	3.17	0.62	<0.000
Prednisolone × ART	-0.83	0.75	0.265	-0.83	0.74	0.264	-0.84	0.74	0.260
Age	-3.12	1.56	0.046	-3.12	1.56	0.045	-3.14	1.60	0.050
Dropout model									
ψ_0	-2.26	-2.26	<0.000	-2.22	0.43	<0.000	-2.07	0.69	0.003
ψ_1	-	-	-	-0.00	0.03	0.910	0.02	0.10	0.826
ω	-	-	-	-	-	-	-0.03	0.126	0.791
-2ℓ		3495.26			3495.24			3495.16	

4.4.2 Results from local influence application to SeM

Figure 4.1 shows the response profiles of CD4 count changes of potentially influential subjects 33, 78, 93, 94 and 133. Subjects with their normal curvatures values greater/less than the corresponding average are considered as influential in the estimation process. These subjects with larger normal curvature values are displayed in Figure 4.2. Figure 4.2 displays the index plots, from the top left, the total influence $C_i(\boldsymbol{\gamma})$, influences for the sub-vectors $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\beta})$, $C_i(\boldsymbol{\alpha})$, $C_i(\boldsymbol{\psi})$ and $\mathbf{h}_{\text{max,i}}$.

Subjects 93 and 94 have larger $C_i(\boldsymbol{\gamma})$ values and are considered as influential for the estimation $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\psi}$. These same subjects have large values for $C_i(\boldsymbol{\psi})$ and $C_i(\tilde{\boldsymbol{a}}_{\text{max,i}})$ and are thus considered

influential for the estimation of $\boldsymbol{\psi}$ and $\tilde{\boldsymbol{a}}_{\text{max,i}}$. Subject 94 has the largest $C_i(\boldsymbol{\gamma})$ value, followed by subject 93. Subject 94's relatively large influence is caused by a large reduction in the subject's CD4 count level 415, after randomization in the placebo arm, to 223 at months 0.5 (2 weeks) visit. The subject 94's relatively large influence is caused by the subject's very weak response to treatment just before dropout. Similar arguments hold for subject 93 with consistently very weak response (43, 81 and 102 at baseline, months 0.5 and 1 respectively) to ART at each scheduled visit before dropout. Subjects 33, 78, and 133 have large values for $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\beta})$ and $C_i(\boldsymbol{\alpha})$ and hence are influential for the estimation of $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$.

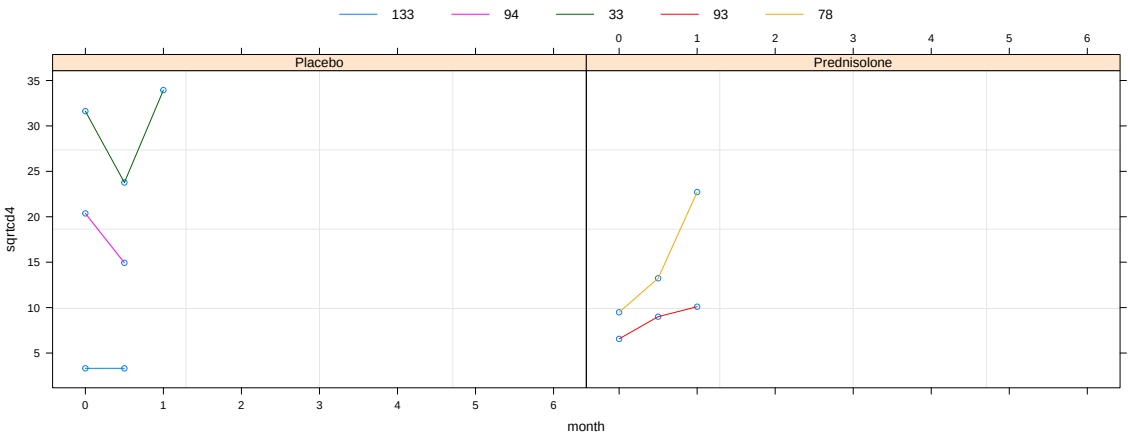

Figure 4.1: Profile plots of influential patients CD4 count evolution over time.

We then considered three analyses without the influential subjects. Firstly, influential subjects 93 and 94 for $\boldsymbol{\gamma}$ were removed; secondly, influential subjects 33, 78 and 133 for $\boldsymbol{\theta}$ were removed; and thirdly, we removed all the influential subjects 33, 78, 93, 94 and 133 from the analyses. We discuss the results of these analyses but display only the results of the first analysis in Table 4.2. The results in the Table 4.2 agree with the results in the Table 4.1. The analyses under the second and third analyses seem to drive the MAR mechanism towards the NMAR mechanism. However, these analyses produced downward-biased estimates of treatment and time effects.

4.4.3 Results from the stress-testing application to the SeM

For the stress-testing methodology (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013), we first set $\omega_{\text{p}} = \omega_{\text{a}} = 0$ (which assumes MAR). To determine the plausible range and combinations of values for ω_{p} and ω_{a}, we estimate $\omega_{\text{p}} = -1.25$ and $\omega_{\text{a}} = -0.79$ under NMAR. Using these estimates, we then consider various combinations of values $(-1.25, -0.79, 0.0,$ and $0.80)$, within the range $-1.25 - 0.80$ for ω_{p} and ω_{a}, in order to assess how estimates of treatment effects will differ for those combinations. Figure 4.3 presents treat-

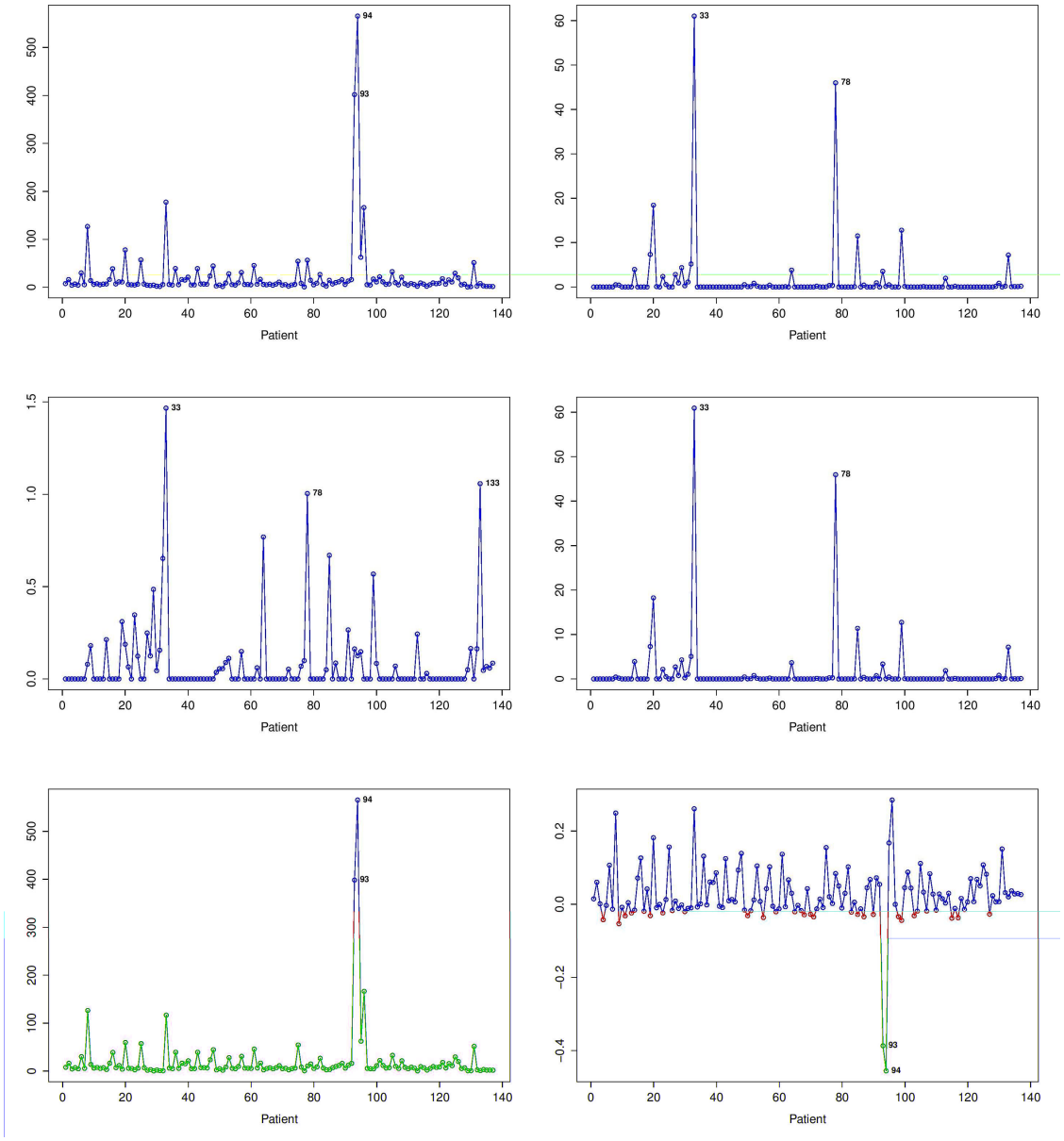

Figure 4.2: Index plots C_i, $C_i(\boldsymbol{\theta})$, $C_i(\boldsymbol{\psi})$, $C_i(\boldsymbol{\beta})$, $C_i(\boldsymbol{\alpha})$ and $\tilde{\boldsymbol{a}}_{\max,\mathrm{i}}$ of maximal curvature.

ment effects and confidence intervals for various combinations of dropout rates from the placebo and prednisolone arms. It can be observed most of 95% confidence intervals for the treatment effects crossed the horizontal (zero line). This gives an indication that the study did not observe a statistically significant difference between the prednisolone and placebo groups. However, significant treatment effects are observed at the $(\omega_p = 0.79, \omega_a = -1.25)$, $(\omega_p = 0.79, \omega_a = -0.79)$, and $(\omega_p = -1.25, \omega_a = 0.79)$ combinations of dropout rates. It can be observed that dropout rates $(\omega_p = 0.79, \omega_a = -1.25)$ and $(\omega_p = 0.79, \omega_a = -0.79)$ show significant increases in the treatment effect whereas $(\omega_p = -1.25, \omega_a = 0.79)$ shows a significant reduction in the treatment effect. These treatment effects are significant under assumed dropout rates outside the range of $-1.25 - 0.79$. Such treatment effects may have clinical interpretation or may be clinically meaningless. For instance, these treatment effects give an indication that the treatment effect in the IMPI trial would have been significant if we could ensure that the dropout probability caused by a positive change on the current outcome values decreases in the prednisolone arm and increases in the placebo arm.

Table 4.2: Parameter estimates for different missingness models fitted to the CD4 count data with subjects 94 and 93 removed.

	MCAR			MAR			NMAR		
Parameter	Est.	s.e	p-value	Est.	s.e	p-value	Est.	s.e	p-value
Measurement model									
Time	0.27	0.10	0.008	0.27	0.10	0.007	0.24	0.13	0.052
Prednisolone	0.45	0.90	0.615	0.45	0.90	0.614	0.39	0.90	0.664
Prednisolone × Time	-0.09	0.15	0.550	-0.09	0.15	0.549	-0.09	0.15	0.516
ART	3.11	0.57	<0.000	3.11	0.57	<0.000	3.22	0.62	< 0.000
Prednisolone × ART	-0.79	0.75	0.293	-0.80	0.75	0.290	-0.80	0.75	0.286
Age	-3.11	1.60	0.052	-3.09	1.58	0.049	-3.16	1.59	0.467
Dropout model									
ψ_0	-2.30	0.16	<0.000	-2.30	0.44	<0.000	-2.05	0.71	0.004
ψ_1	-	-	-	0.000	0.03	0.999	0.04	0.11	0.698
ω	-	-	-	-	-	-	-0.06	0.14	0.686
-2ℓ		3454.36			3454.36			3454.14	

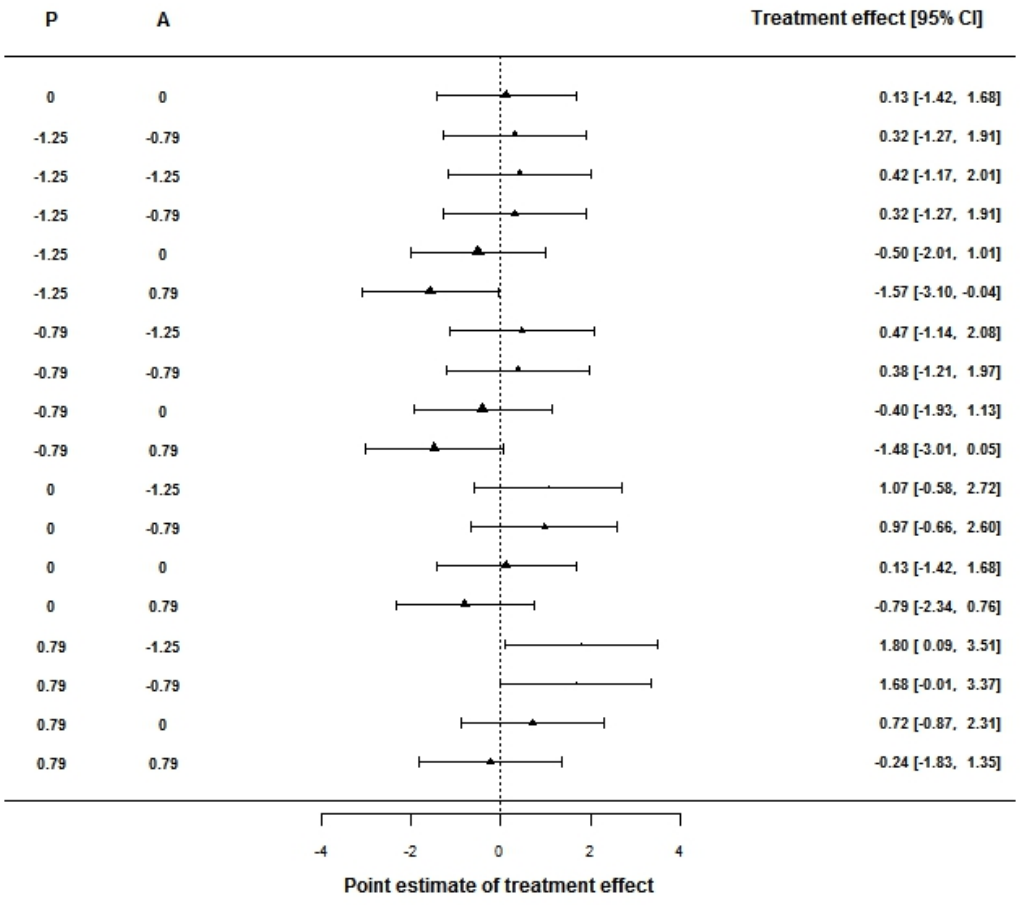

Figure 4.3: Forest plot of 95% CI of treatment effects for different combinations of dropout rates at the placebo (**P**: ω_p) and prednisolone (**A**: ω_a) arms.

4.5 Discussion

In this chapter, we have investigated the effect of prednisolone and ART treatments on CD4 changes over time and seek to draw statistical inferences in the presence of missing data by conducting sensitivity analyses to the missing CD4 count measurements from the IMPI trial. The sensitivity analysis are based on the Diggle and Kenward (1994) selection model framework and its extension to the local influence (Verbeke et al., 2001) and stress-testing (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013) approaches.

The study results, from these sensitivity approaches, showed that there is significant linear time trend and that there is no significant prednisolone effect. Prednisolone does not interact with ART treatment; subjects who received ART have higher CD4 count levels, and older subjects had lower CD4 count levels. The results also showed that inferences under the MAR model are robust to the analyses under the NMAR model and that influential subjects do not drive the MAR mechanism towards the NMAR mechanism.

The results under the NMAR models for sensitivity analysis (local influence and stress-testing) are comparable to the results under the primary MAR model. These findings are in line with Molenberghs et al. (2008) paper in a sense that "for every MNAR model fitted to a set of incomplete data, can be replaced by an MAR version which produces exactly the same fit to the observed data (Molenberghs et al., 2008)". Since our results showed that the statistical inferences under the MAR model are robust to statistical inferences under the NMAR models considered, one can make statistical inferences using the results obtained under the MAR model. Molenberghs et al. (2008) noted that unless one puts strong *a priori* belief in the posited MNAR model, it is not possible to use the fit of an MNAR model for or against MAR. Under the local influence, a downward-biased estimates of the treatment and time effects are observed in analyses where influential subjects drive the MAR mechanism towards the NMAR mechanism. The results from the stress-testing showed that the treatment effect in the IMPI trial would have been significant if the dropout probability caused by a positive change on the current outcome values decreases in the prednisolone arm and increases in the placebo arm, which is unlikely.

Our study is intended to assess the importance of conducting some forms of sensitivity analysis and to illustrate these principles in the IMPI trial setting. The IMPI trial was a cardiology trial in which HIV data were relevant and collected. However, the HIV data were not collected as they would have been in a HIV-related clinical trial, and hence there are missing CD4 counts. Although the IMPI trial was a cardiology trial, our analyses of the HIV data provide reasonable information regarding the degree of influence of ART and prednisolone treatments on CD4 changes over time. The missingness of CD4 values might be informative, and hence later values of CD4 counts might

be missing because subjects have died. This would require joint modeling on the CD4 counts and time to death.

The DK SeM methodology considered in this study is used to assess the effect of influential subjects on parameter estimates and conclusions (Thijs et al., 2000; Verbeke et al., 2001). However, the methodology can also be used to investigate the influence of individual observations. In addition, there are other methods of identifying outliers/influential observations which could be used within the DK model. Examples of such methods are the variance shift outlier model (VSOM) (Gumedze et al., 2010) and Cook's distance (Cook, 1986). This is part of on-going research work.

The sensitivity analyses approaches based on the DK SeM (Diggle and Kenward, 1994) were applied to monotone data. However, these sensitivity analyses approaches can be applied to non-monotone data (Molenberghs et al., 2008). Also, the sensitivity analyses approaches were applied to continuous longitudinal measurements. However, the approaches can also be applied in incomplete longitudinal ordinal data settings (Van Steen et al., 2001). Jansen et al. (2006) applied local influence approach to binary data obtained from a psychiatric study. Related methods for generalized linear models with missing covariates were proposed by Shi et al. (2009) and Rakhmawati et al. (2016, 2017) proposed local influence approach in the context of the generalized linear mixed effect model to handle incomplete over-dispersed count data as well as binary and time-to-event data (Rakhmawati et al., 2017).

It is worth noting that the methods used in this thesis can also be applied to other models. For instance, the methods can be used to fit models with visit as a class variable and with the heterogeneous first-order autoregressive covariance structure Shen et al. (2006). Others covariance structures incorporating the random effects and serial correlations can be used (Michiels et al., 2002). The approach can also be applied to models where compound symmetry covariance structured can be assumed. The local influence approach is able to detect unusual features of study subjects that may not necessarily be related to the dropout mechanism (Verbeke et al., 2001).

Verbeke et al. (2001) found that subjects with an unusually high profile, or a somewhat atypical serial correlation behavior, are detected with the local influence tool. At first sight, this raises a concern because the ω_i parameter is specified in the dropout model and not in the measurement model. This calls for further investigation regarding which effects are easy or difficult to detect using the local influence methods. Jansen et al. (2006) further studied the behavior of sensitivity assessment tools based on the effect of sample size, effect of anomalies in the missingness mechanism, effect of anomalies in the measurement model and then noted that the methodology, while useful, should be applied with caution. These authors noted that the occurrence of influential subjects calls for further investigation, but there are no automated rules as to what a particular deviation precisely means.

In this chapter, we have made contribution to research by conducting sensitivity analyses to show how far dropout probability would have been in order to overturn treatment effect. These analyses would allow medical and clinical researchers to stress-test assumptions about the missing data and also to assess the effect of such assumptions on the statistical inferences. Such analyses also help the researchers to offer clinical explanations associated with the results obtained. For instance, we found that treatment effect in the IMPI trial would have been significant if we could ensure that the dropout probability caused by a positive change on the current outcome values decreases in the prednisolone arm and increases in the placebo arm. This finding is clinically meaningful in the context of the IMPI trial in a sense that a reduction in the amount of unobserved CD4 count data in treatment arms is likely to yield significant treatment effect. Clinically, this may imply that with higher retention rate in the prednisolone arm, the effect of interaction between ART and prednisolone is likely to be significant leading to an increase ART effect (with higher CD4 count values) for patients in the prednisolone arm relative to the prednisolone placebo arm. The results from these approaches as well as a SAS code for their implementation have been published in (Iddrisu and Gumedze, 2018) and also attached in the Appendix A.

Chapter 5

Pattern-mixture model with multiple imputation

5.1 Introduction

In this chapter, we introduce an alternative sensitivity approach to missing data. This sensitivity analysis approach is known as the pattern-mixture with multiple imputation (PM-MI) (Carpenter et al., 2013) and is based on the standard pattern-mixture model (PMM) introduced in the Chapter 3, Section 3.1.2.

We will first define the estimands and their associated deviations as well as some key assumptions relevant for the sensitivity approaches in this chapter. Thereafter, we introduce the conventional PMM and then focus on the PM-MI methodology (Carpenter et al., 2013). This is followed by a brief discussion of the assumptions (sensitivity analysis) that allow us to obtain missing post-deviation data under the PM-MI approach.

Carpenter et al. (2013) applied the PM-MI approach to measurements at the last visit, where a linear regression model for the final observation on baseline and treatment is fitted to each of the imputed dataset. In this chapter, we considered the PM-MI approach for incomplete longitudinal, where we assumed the linear mixed model (Laird and Ware, 1982) for the observed longitudinal measurements. The PM-MI approach is then applied to simulated datasets and the incomplete longitudinal CD4 count data from the IMPI trial and the results discussed.

5.2 Estimands for primary and sensitivity analyses

This section presents the *de jure* and *de facto* estimands (Carpenter et al., 2013). This discussion is necessary because our primary analysis model is based on the *de jure* estimand, and the sensitivity analysis models are based on the *de facto* estimand (Carpenter et al., 2013). The primary analysis (as specified in the statistical analysis plan) addresses the main objective of the study, whereas the sensitivity analysis considers models that make alternative assumptions (trial protocol) that, in one way or another, may influence statistical inferences under the primary analysis model. We discuss the *de jure* estimand in Section 5.2.1 and then the *de facto* estimand in Section 5.2.2. We will also discuss deviations associated with each estimand in Section 5.2.3.

5.2.1 De jure estimand hypothesis

The *de jure* estimand estimates the effect of treatment on patients assuming that patients adhered to the study protocol without deviating from the trial protocol (Carpenter et al., 2013; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Diggle and Kenward, 1994; Creemers et al., 2011). The *de jure* estimand hypothesis may be defined in terms of the MAR mechanism as the conditional distribution of observations later in the follow-up, given observations earlier in the follow-up, is independent of whether deviation occurs. In this case, patients are expected to obtain the full benefit of the treatment and the question of interest is whether the treatment works under the best-case scenario. The *de jure* estimand hypothesis can also be defined as that patients with missing values continued on treatment and got a greater improvement than those whose data are observed.

In this study, the *de jure* primary analysis is based on the multiple imputation under missing at random (MAR) (Rubin, 1996, 1976, 1987). The primary analysis method to choose varies from trial to trial. The guidelines on how to decide on an appropriate primary analysis for a given trial can be found in the NRC panel report (NRC, 2010) and many others (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Permutt, 2016).

5.2.2 De facto estimand hypothesis

The *de facto* estimand concerns what would be the effect of treatment seen in practice if treatment were allocated to the target population of eligible patients as defined by the trial inclusion criteria. In addressing this question, we may ask, what would have been the effect of treatment seen at the end of the study if those who deviated moved to the equivalent of the active treatment arm (prednisolone treatment in this study). However, this may underestimate the benefit of active treatment in trials where more benefit is expected from the active treatment. This is because estimand equates treatment benefit of those failing on the placebo arm to those opting for active treatment. In this instance, the fairer comparison might be to move those who deviate from the prednisolone arm onto the placebo arm. In the case of the IMPI trial, since all patients in both prednisolone and placebo arms were given ART, we expect no significant difference in their response to ART treatment unless there is interaction between prednisolone and ART treatment.

We discuss four *de facto* options for obtaining post-deviation data in Section 5.4. These options make assumptions about the missing post-deviation data. These assumptions are alternative plausible assumptions, which depart from the MAR assumption under the primary analysis. In this way, it is assumed that the data are not missing at random (NMAR) and we assess the robustness of inferences under the MAR primary analysis to the alternative assumptions under the *de facto*

options (sensitivity analyses).

5.2.3 Deviations associated with estimands

It is important to define clearly deviations associated with each estimand in the study protocol. This is because clarity of deviations associated with each estimand is vital for primary analysis and framing relevant sensitivity analysis (Carpenter et al., 2013). The exact definition of a deviation will depend on the trial setting and may also vary between separate analyses (Carpenter et al., 2013). In the IMPI trial, the following situations can be regarded as deviations associated with the *de jure* estimand: unblinding of treatment arms and unobserved CD4 count measurements and deviations associated with the *de facto* estimand are unblinding such as treatment allocation, loss to follow-up such that no further treatment is taken and influence if trial prednisolone treatment on ART.

Given the estimands and their associated deviations, it is assumed that each patient has longitudinal follow-up data until either the patient deviates or reaches the final visit, and that the nature or reason of each deviation is known. This approach further assumes that for each deviation or group of similar deviations occurring in a dataset due to similar reasons, an appropriate post-deviation distribution can be built taking into consideration (1) the patient's pre-deviations, (2) pre-deviations and post-deviations data from other patients in the trial, (3) the nature of the deviation, and (4) the reason for the deviation (Carpenter et al., 2013). In the IMPI trial CD4 count data, the nature of the deviation was that the centres did not have enough resources to measure CD4 count data for all patients leading to unobserved CD4 count data.

5.3 Standard pattern-mixture model and the pattern-mixture model with multiple imputation

It this section, we give a brief review of the standard pattern-mixture model (PMM) and then discuss the pattern-mixture model with multiple imputation (PM-MI) (Carpenter et al., 2013). In Section 5.3.3, we give the link between these approaches.

5.3.1 Standard pattern-mixture model

We have mentioned in the Chapter 3 that the pattern-mixture modeling framework is a reverse factorization of the selection model (Heckman, 1976; Rubin, 1976; Diggle and Kenward, 1994; Little and Yau, 1996). The PMM approach (Wu and Bailey, 1988; Wu and Carroll, 1988; Wu and Bailey,

1989) is defined as

$$\begin{aligned}\Pr\left(\mathbf{Y}_i, \mathbf{R}_i \mid \mathbf{X}_i, \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \mathbf{R}_i, \boldsymbol{\theta}\right) \\ \times \Pr\left(\mathbf{R}_i \mid \mathbf{X}_i, \boldsymbol{\psi}\right).\end{aligned} \tag{5.1}$$

The PMM has desirable properties especially where the data are NMAR. For instance, where it is not substantively reasonable to consider non-responses as missing data, it may be desirable to limit the inferences to the subpopulation of patients whose responses are observed (Carpenter et al., 2013). Thus, it is more meaningful to consider the distribution of $\mathbf{Y}_i$ given $\mathbf{R}_i = 1$ rather than the marginal distribution of $\mathbf{Y}_i$ (Rubin, 1976). Contrary to the selection model, $\Pr\left(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{X}_i, \mathbf{R}_i\right)$ is modeled directly from the pattern-mixture model.

One important feature of the pattern-mixture model (5.1) is that it fits a different response model for each pattern of missingness such that the observed data is a mixture of patterns weighted by their respective probabilities of missing patterns. That is, the first component in the PMM (5.1), $\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \mathbf{R}_i, \boldsymbol{\theta}\right)$ fits a response model for each pattern of missingness and $\Pr\left(\mathbf{R}_i \mid \mathbf{X}_i, \boldsymbol{\psi}\right)$ represents dropout probability for each pattern. It follows that if there are U number of missingness patterns in a data set, then the marginal distribution of $\mathbf{Y}_i$ is a mixture of $\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_i, \boldsymbol{\theta}\right) = \sum_{u=1}^{U} \Pr\left(\mathbf{Y}_i \mid \mathbf{R}_i = \mathbf{R}_i^u, \mathbf{X}_i, \boldsymbol{\theta}^u\right) \pi_u$, where $\pi_u = \Pr\left(\mathbf{R}_i = u \mid \mathbf{X}_i, \boldsymbol{\psi}\right)$ and $\mathbf{R}_i$ counts the number of U patterns, $\boldsymbol{\theta}^u$ represents the parameters of marginal density $\Pr\left(\mathbf{Y}_i\right)$ in the u^{th} pattern. It can be observed that in the pattern-mixture model, parameters $\{\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^U\}$ can have different dimensions. A logistic model is often assumed for dropout probabilities and a linear mixed effects model (2.3) (Laird and Ware, 1982) for the measurement process.

The pattern-mixture model (5.1) is well understood using the second MAR assumption. The second MAR assumption states that observations that would have been recorded for a patient in the future given the observed history of such a patient has the same statistical behavior. This feature of the pattern-mixture model makes it possible for multiple imputation to provide a practical approach to estimation and inferences. In addition, this feature provides a framework for the formulation of the pattern-mixture model with multiple imputation (Carpenter et al., 2013).

5.3.2 Pattern-mixture model with multiple imputation methodology

In this section, we describe the pattern-mixture model with multiple imputation (PM-MI) methodology (Carpenter et al., 2013). Consider a randomized clinical trial with two treatment arms and predictors of continuous response $\mathbf{Y}_i$ (Y_{ij}) for each patient. Let the Y_{ij} be the measurements of the ith patient at the jth occasion in each treatment arm $\mathbf{T}_i$, where $j = 0$ represents baseline measurements in each treatment arm and $j = n$ denotes the last observation time prior to a deviation for the ith patient. It is then assumed that all patients were observed at baseline. Let

(1) $\mathbf{Y}_i^o = (Y_{i0}, \ldots, Y_{in})'$ denote a vector of the ith patient's observed responses at each scheduled visit from $j = 0, \ldots, n$, (2) $\mathbf{Y}_i^m = (Y_{in+1}, \ldots, Y_{in})'$ denote a vector of the ith patient's missing post-deviation responses at scheduled visits time from $j = n_i + 1, \ldots, n$, where n is the last schedule visit, (3) $\mathbf{Y}^m = \left(\mathbf{Y}_1^{m'}, \ldots, \mathbf{Y}_N^{m'}\right)'$ denotes a column vector of the ith patient's missing post-deviation responses profile, and (4) $\mathbf{Y}^o = \left(\mathbf{Y}_1^{o'}, \ldots, \mathbf{Y}_N^{o'}\right)'$ denotes a column vector of the ith patient's observed responses profile. It follows that the distribution of each patient's post-deviation responses $\mathbf{Y}_i^m$, given each patient's pre-deviation responses $\mathbf{Y}_i^o$ and the deviation time n, is defined by

$$\Pr\left(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, n, \mathbf{T}_i, \boldsymbol{\theta}\right), \tag{5.2}$$

where $\mathbf{T}_i$ denotes binary treatment arm (for patient in either the prednisolone or placebo treated arm). The parameter vector $\boldsymbol{\theta}$ has to be estimated before we can impute missing post-deviation data by drawing from conditional distribution (5.2).

5.3.3 Link between the pattern-mixture model and the pattern-mixture model with multiple imputation

If post-deviation data are assumed to be MAR (that is, the probability that the responses are missing depends on the observed data), the distribution (5.2) is independent of the deviation time n. Hence the distribution (5.2) can be written as

$$\Pr\left(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{T}_i, \boldsymbol{\theta}\right). \tag{5.3}$$

Under such an assumption, the direct maximum likelihood estimation (Rubin, 1976; Dempster et al., 1977) or the multiple imputation under MAR can be used to obtain valid inferences (Rubin, 1976, 1996, 1977). However, if data are NMAR, the distribution (5.2) depends on the deviation the time n in a manner that could be different for each patient. This feature of the distribution (5.2) is analogous to the standard pattern-mixture model (5.1), where the response model is fitted for each pattern of missingness such that the observed data is a mixture of patterns weighted by their respective probabilities of missingness.

It follows that for each patient or group of patients, a specific form of the conditional distribution (5.2) is defined to reflect a specific assumption appropriate to their treatment arm $\mathbf{T}_i$, deviation time n_i and other relevant information or covariates. Given this information, multiple imputation is used for imputing missing post-deviation data from equation (5.2) to create complete data sets. Thereafter, estimation and inference is then performed by fitting a standard method of analysis (which is a methods of analysis that yields valid inferences without missing data) to the complete data sets (Rubin, 1996; Little and Rubin, 1987). Thus, for inferences about $\boldsymbol{\theta}$ in the presence of deviations, multiple imputation is used to create K "completed" data sets.

To obtain post-deviation data from the distribution (5.2), Carpenter et al. (2013) suggested the following.

Step A: Assume a multivariate normal for the observed data $\mathbf{Y}^o$.

Step B: Draw samples of the parameter estimates of $\boldsymbol{\beta}$ and $\mathbf{R}_i$ from the Bayesian posterior distribution defined as $\Pr(\boldsymbol{\beta}\prime, \boldsymbol{\alpha}\prime \mid \mathbf{Y}^o)$, where $\boldsymbol{\beta}$ is a vector of the means and $\boldsymbol{\alpha}\prime = (\boldsymbol{\sigma}\prime, \boldsymbol{\rho}\prime)\prime$ is a parameter vector of the variance components in the measurement model. The Markov chain Monte Carlo (MCMC) method is used to draw samples of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ from this posterior.

Step C: Update the Markov chain sufficiently after each draw in order to avoid correlation between draws in each of the parameter estimates $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$.

Step D: After each draw of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ for each patient who deviates before the end of the trial, $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are used to build the joint distribution of such patient's pre-deviation and post-deviation data. We discuss different options for building this joint distribution in Section 5.4.

Step E: The joint model in Step **D** is then used to build the conditional distribution of each patient's missing post-deviation data, given the pre-deviation data (5.2). The missing post-deviation data in the conditional distribution (5.2) are obtained using the parameter estimates $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ obtained from **Step D**.

Step F: Repeat Steps **B-E** K times to create K "complete" data sets. Thereafter, any method of analysis that yields valid inferences in the absence of missing data can then be applied to the complete data sets.

Carpenter et al. (2013) considered the treatment benefit at the last scheduled visit where they fitted a linear regression model that assumed that observations are independent. This thesis considers the treatment benefit over time and hence the linear mixed effect model (Laird and Ware, 1982) is assumed for the measurement process. This model is then fitted to each of the K imputed data sets. This analysis produced K statistics for the parameter estimates $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$. Estimates from each of the K completed data set were then combined to produce single estimates with their associated standard errors using the Rubin's rule (Rubin, 1996).

5.4 Constructing joint distributions of pre-deviation and post-deviation outcome data

In this section, we discuss the four *de facto* options for obtaining the missing post-deviation data (Carpenter et al., 2013). These options make alternative and plausible assumptions about the missing data such that the *de facto* (NMAR sensitivity analysis) assumptions depart from the

de jure (MAR primary) assumption about the missing data. These assumptions assess whether inferences under such MAR primary analysis assumption are sensitive to the alternative plausible assumptions under NMAR sensitivity analysis. In this way, we will be able to assess whether the process that generated the missing CD4 count data is MAR or NMAR mechanism. This distinction is necessary because the type of missing data mechanism has implications for both the analysis and interpretations of the results (Molenberghs et al., 2008). We also discuss how to choose reference arm (Section 5.4.1) and the implications of the *de facto* options under the IMPI trial in Section 5.4.2.

Carpenter and colleagues proposed the following options for constructing the joint distribution of each patient's pre- and post-deviation outcome data where each option represents a possible *de jure* or *de facto* assumption concerning post-deviation data. These assumptions differ in the ways in which unavailable information for deviated patient are borrowed, or estimated, from other groups of patients in the same trial (Carpenter et al., 2013). Here two treatment arms, placebo and active (prednisolone in our study), are considered and one of these arms is chosen as a reference arm such that unavailable information for the deviated patient can be "borrowed" from such reference arm. The reference arm could be either the placebo or the active arm depending the hypothesis to address. In this study, we in turn used each arm as a reference arm just to explore how treatment effect is affected under such considerations. Here, we refer to the arm not chosen as reference as the other arm.

A: *Jump to reference (J2R)*: Under this assumption, after a patient stops taking treatment from the randomized arm, such patient's mean response distribution is now considered to be the same as that of the "reference" group of patients. Typically, such a patient will take treatment from the control or placebo arm. However, such a patient may not necessarily take treatment from the placebo arm (but assumes to take treatment from the randomized arm after dropout) since the choice of the reference arm may depends on trial setting. In a trial where more benefit is expected in the active arm, such a change may be seen as extreme, and choosing the reference group to be the placebo group may be viewed as a worst-case scenario in terms of reducing any treatment benefit, since withdrawn patients on active will lose the effect of their period on treatment. In this study, the post-deviation data in the reference arm are imputed under randomized-arm MAR.

B: *Copy difference in reference (CDR)*: Under this *de facto* option, after the patient deviates, it is assumed that the patient's post-deviation mean increments copy those from the reference arm. For instance, if the placebo arm is chosen as the reference arm, the patient's mean profile after deviation tracks that of the mean profile in the placebo arm, but starts from the benefit already obtained from the active arm.

C: *Last mean carried forward (LMCF)*: Under the LMCF, it is assumed that after deviation, the

patient's post-deviation means equal that of the marginal mean of the randomized treatment arm.

D: *Copy reference (CR)*: The "copy reference" de facto option assumes that a patient's whole distribution, both pre-deviation and post-deviation data, is the same as reference arm.

Whereas the above assumptions for constructing post-deviation data have been proven to be practical and permit relevant, accessible assumptions for framing primary and sensitivity analyses, the PM-MI approach depends on the relevance of the assumptions about missing post-deviation data in relation to the context of the trial at hand (Carpenter et al., 2013). In this study, we apply the PM-MI approach in the context of the IMPI trial setting (see Sections 5.4.1 and 5.4.2).

5.4.1 Choosing the reference arm

For the "jump to reference", "copy reference" and "copy increment in reference" *de facto* options, we discuss the implications for the choice of the reference arm. In the IMPI trial, it could be either the placebo or the prednisolone arm. This is because we expect similar statistical behaviour for patients in either arm. We expect similar statistical behaviour because patients in both the prednisolone and the placebo arms were given ART at randomization or when convert to HIV+. Suppose that one wishes to address the *de facto* question corresponding to the assumption that after post-deviation (CD4 count measurements are unobserved), (1) patients on the placebo arm obtain a treatment equivalent to the active (prednisolone) arm, and (2) the prednisolone-treated patients continue on treatment and adhered to the study protocol, so that their post-deviation data can be imputed assuming randomized-arm MAR. In such a case, we specify the prednisolone arm as a reference.

In the IMPI trial, HIV+ patients in either placebo or prednisolone arm were given ART and thus patients with their CD4 count unobserved are expected to have equivalent treatment benefit compared with those patients with their CD4 count observed unless prednisolone treatment influences ART treatment. Since we hypothesized that patients' response to ART treatment in both the placebo and the prednisolone arms are comparable, we also present results where the placebo arm is used as a "reference". Thus dropouts in the prednisolone arms obtain treatment equivalent to the placebo arm so that their post-deviation data (unobserved CD4 count measurement) can be imputed under randomized-arm MAR. This latter assumption might be appropriate where no alternative treatment is generally available or where patients in both arms receive treatment but responses were unobserved (in the case of the IMPI trial IMPI trial).

5.4.2 De facto options under the IMPI trial

A simple interpretation of the PM-MI approach is that within the same trial, the PM-MI approach is used to "borrow" or estimate unavailable information from a group of patients for another group of patients who have their information missing. As we have stated earlier, in the IMPI trial setting, HIV+ patients in both the active treatment (prednisolone) arm and the placebo treatment arm were given ART, and hence we expect similar benefit from ART treatment unless prednisolone treatment interacts with the ART treatment.

One research question to address in the IMPI trial is whether the prednisolone treatment interacts with the ART treatment. If they do interact, patients' response to ART treatment from the active arm and the placebo arm will be different, otherwise they would be comparable. Also in the IMPI trial, missing CD4 counts for patients were unobserved due to inadequate resources but not necessarily that the patient dropped out before the end period of the trial. In other words, CD4 count measurements were missing at some scheduled visits mostly due to administrative reasons and missingness would have been generated by a random process.

In fact, only 6% of the patients dropped out (genuine dropout) in the IMPI trial. This means that most of the patients do not dropout (but continued to receive treatment) from the study but their CD4 count values could not be measured due to inadequate resources. Thus, patients who CD4 counts are unobserved, are expected to have similar CD4 count levels to those who were observed. Out of a total number of 294 HIV positive patients in the placebo arm, approximately 78% were already on ART at the time of randomization and out of a total number of 293 HIV positive patients in the prednisolone arm, approximately 80% were already on ART at baseline.

For the *de facto* question, since we do not expect significant different in treatment effect between patients with their CD4 count observed and those with their CD4 count unobserved, the jump to reference and the copy reference options are the most plausible options for assessing sensitivity of inferences to MAR assumption.

The CD4 count data introduced in Section 2.6, are analyzed under *de jure* MAR and *de facto* NMAR assumptions. The linear mixed effects model (4.6) is fitted to the CD4 count measurement.

5.5 Analyses of the CD4 count data

This section presents PM-MI analyses of the monotone and non-monotone CD4 count data. We implemented the PM-MI approach using STATA **mimix** package developed by Cro of London School of Hygiene and Tropical Medicine (LHTM), UK. This package imputes missing continuous

outcomes for a longitudinal trial with protocol deviations under distinct reference groups based assumptions for the unobserved data, following the procedure proposed by Carpenter et al. (2013).

To address the *de jure* hypothesis, we performed multiple imputation for the unobserved CD4 count under MAR mechanism using the **ice** package in STATA (Royston, 2005). We also impute post-deviation under LMCF, J2R, CDR and CR *de facto* options to obtain a complete data sets. The linear mixed effect model (4.6) was then fitted to each of the completed data sets and parameter estimates combined to produce parameter estimates with their corresponding standard errors using the Rubin's rule (Rubin, 1996; Royston, 2005).

5.5.1 Analyses of the monotone CD4 count data

In this section, our analysis is for monotone pattern. We consider the jump to reference option for illustration purpose and Figure 5.1 shows profiles plots of the mean $\sqrt{\text{CD4}}$ count measurements for the complete data sets, for each deviation pattern, by placebo arm (Treatment = 0) and prednisolone arm (Treatment = 1).

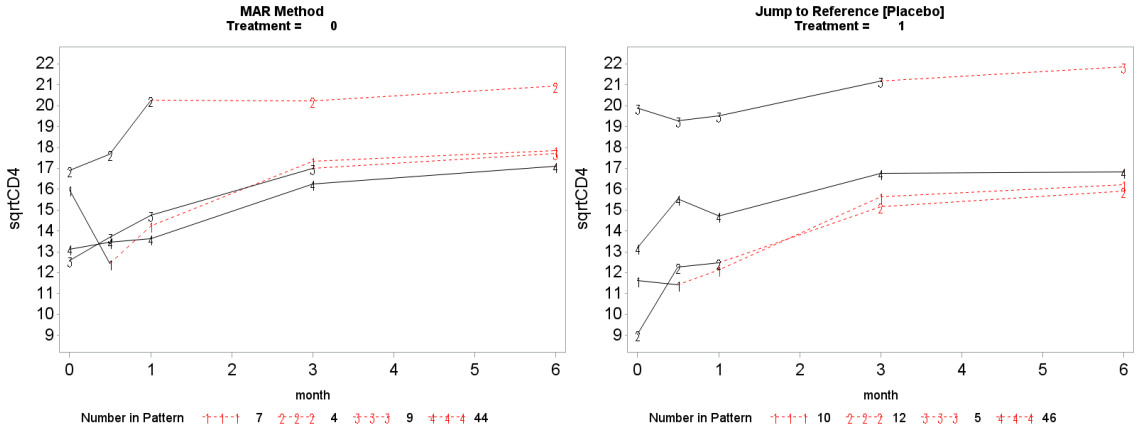

Figure 5.1: Placebo reference arm (Treatment = 0): Profile plots of the mean $\sqrt{\text{CD4}}$ count against month for the four different deviation patterns. The solid lines join the observed means (before deviation) and the dotted lines join the means of the imputed data for that pattern. Pattern 4: group of patients who completed the study (completers); Pattern 3: group of patients who dropped out after month 3; Pattern 2: group of patients who dropped out after month 1; and Pattern 1: group of patients who dropped out after week 2.

The left panel of the Figure 5.1 shows complete data profiles of the placebo reference arm with missing post-deviation values obtained under MAR, whereas the right panel of the Figure 5.1 shows complete data profiles of the prednisolone arm patients with missing post-deviation data "borrowed" from the placebo arm (left panel of Figure 5.1). We in turn used the prednisolone arm as a reference where the complete data profiles are shown in Figure 5.2. The right panel of the Figure 5.2 shows complete data profiles of the prednisolone reference arm with missing post-deviation values obtained under MAR, whereas the left panel of the Figure 5.2 shows complete data profiles of the placebo

arm patients with missing post-deviation data obtained from the prednisolone arm (right panel of the Figure 5.2). It can be observed that treatment seems to reduce the CD4 count a little, and so imputed data for placebo under MAR are above those in which the placebo patient jumps to the prednisolone arm. Therefore, we investigate the significance of such reduction in the CD4 count level by using the parameter estimates associated with the prednisolone-ART interaction (see Table 5.1).

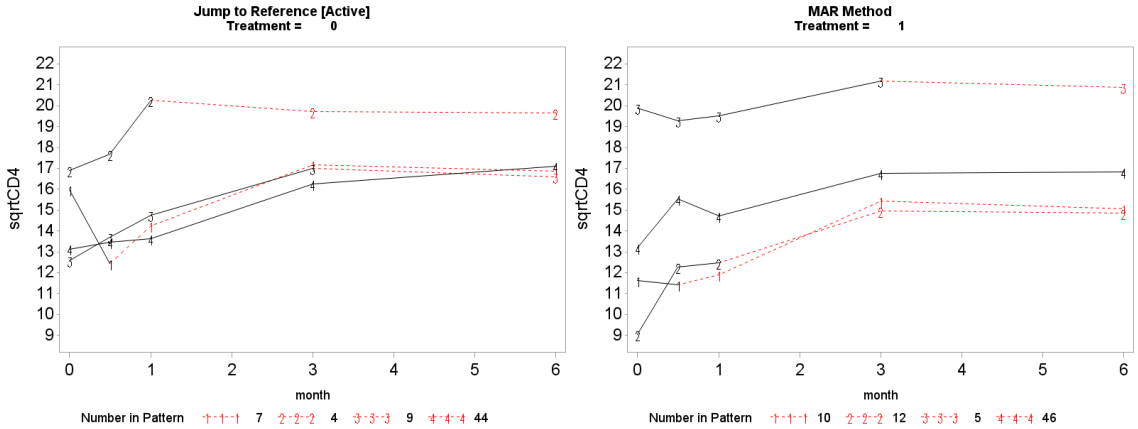

Figure 5.2: Prednisolone reference arm (Treatment = 1): Profile plots of the mean $\sqrt{\text{CD4}}$ count against month for the four different deviation patterns. The solid lines join the observed means (before deviation) and the dotted lines join the means of the imputed data for that pattern. Pattern 4: group of patients who completed the study (completers); Pattern 3: group of patients who dropped out after month 3; Pattern 2: group of patients who dropped out after month 1; and Pattern 1: group of patients who dropped out after week 2.

Similar plots for LMCF, CDR and CR can be found in **Appendix** B.1. After imputation of the missing post-deviation data under LMCF, J2R, CDR and CR, we fit a linear mixed effect model (4.6) to the completed data sets and combine the parameter estimates from each data set using the Rubin's rule to produce parameter estimates with their associated standard errors for the final inferences. The parameter estimates from these analyses are shown in Table 5.1.

Table 5.1 shows that inferences from the MAR primary analysis (MI), which addresses the *de jure* hypothesis, are robust to the difference assumptions under the NMAR sensitivity analyses under *de facto* estimand hypothesis (LMCF, J2R, CR, and CDR). This result thus serves as a justification that the mechanism that generated the missing data in the CD4 count measurements from the IMPI trial is missing at random (MAR) mechanism. The implication of this justification is that the direct maximum likelihood and multiplication methods under MAR can be used to provide valid inferences when assessing the effect of prednisolone and ART treatments on changes in CD4 count level among different treatment groups.

The results show that there is no significant prednisolone effect. The effect of prednisolone-ART is also not significant. This confirmed our hypothesis that prednisolone treatment does not influence

Table 5.1: Monotone pattern: Parameter estimates from *de jure* and *de facto* analyses (N = 137 patients)

	Prednisolone			**Time**			**Prednisolone x Time**			**ART**			**Prednisolone x ART**			**Age**		
Analysis[a]	Est	s.e.	p-value	Est	s.e.	p-value	Est.	s.e.	p-value	Est.	s.e.	p-value	Est.	s.e.	p-value	Est.	s.e.	p-value
MI	0.56	0.791	0.481	0.39	0.140	0.005	0.13	0.195	0.494	2.76	0.689	<0.0001	-1.12	0.966	0.245	-3.12	1.372	0.023
LMCF	0.46	0.870	0.597	0.34	0.092	<0.0001	-0.05	0.146	0.751	2.81	0.548	<0.0001	-0.75	0.875	0.400	-3.06	1.529	0.046
J2R (P-)	0.42	0.845	0.619	0.53	0.094	<0.0001	0.04	0.161	0.825	2.33	0.556	<0.0001	-0.96	0.842	0.260	-3.00	1.509	0.047
J2R (P+)	0.42	0.849	0.623	0.42	0.103	<0.0001	0.02	0.130	0.879	2.57	0.544	<0.0001	-0.96	0.786	0.275	-3.00	1.518	0.048
CDR (P-)	0.46	0.851	0.586	0.52	0.094	<0.0001	0.05	0.148	0.729	2.33	0.56	<0.0001	-0.96	0.844	0.259	-2.97	1.507	0.049
CDR (P+)	0.46	0.863	0.593	0.41	0.113	0.001	0.027	0.147	0.852	2.59	0.589	<0.0001	-0.88	0.882	0.3230	-2.98	1.52	0.050
CR (P-)	0.43	0.852	0.613	0.53	0.093	<0.0001	0.002	0.144	0.990	2.34	0.554	<0.0001	-0.87	0.815	0.288	-3.00	1.509	0.047
CR (P+)	0.44	0.858	0.608	0.43	0.114	<0.0001	0.006	0.137	0.966	2.55	0.579	<0.0001	-0.839	0.848	0.328	-3.00	1.521	0.049

[a]MI: multiple imputation; LMCF: last mean carried forward; J2R (P-): jump to placebo reference arm; J2R (P+): jump to active reference arm, CDR (P-): copy difference with placebo reference arm; CDR (P+): copy difference with active reference arm; CR (P-): copy reference with placebo reference arm; and CR (P+): copy reference with active reference arm

ART treatment. However, there seems to be a slight reduction of CD4 count level in the prednisolone arm. Patients' CD4 count levels increased significantly with time, and patients who are permanently on ART have significantly higher CD4 count levels relative to those who were not ever on ART treatment. The prednisolone-time interaction results show a very slight increase in CD4 count level in the placebo arm compared with prednisolone arm over time. However, this increase is not significant. The near-zero estimates of the prednisolone-time interaction effect suggest that there is no difference in prednisolone effect in both arms over time. This means that the effect of treatments in both arms does not differ significantly over time.

The results also show that older patients are more likely to have a lower CD4 count, hence their CD4 count levels significantly decreases with increasing age. These results agree with the mean $\sqrt{\text{CD4}}$ count profiles plots in Figure 2.2 and Figure 2.4. This is because CD4 count in both the prednisolone and placebo arms increases at the same rate (no significant prednisolone effect and prednisolone does not influence ART treatment) and CD4 count increases with increasing time where this increase, in both arms, is the same over time (no prednisolone-time effect).

5.5.2 Analyses of the non-monotone CD4 count data

This section presents the PM-MI analyses of the combined monotone and non-monotone data. Parameter estimates of these analyses are shown in Table 5.2. The results of these analyses agree with the results under Table 5.1. These results also give an indication that the MAR primary analysis (MI), which addresses the *de jure* hypothesis, are robust to the difference assumptions by and the NMAR sensitivity analyses under *de facto* estimand hypothesis (LMCF, J2R, CR, and CDR). These analyses show that the mechanism that generated the missing data in the CD4 count measurements from the IMPI trial is the missing at random (MAR) mechanism. This means that the direct maximum likelihood and multiplication methods under MAR can be used to provide valid inferences when assessing the effect of prednisolone and ART treatments on changes in CD4 count level among different treatment groups.

It can be observed from these analyses that there is no significant prednisolone effect and the effect of prednisolone-ART is also not significant. This implies that prednisolone treatment does not influence ART treatment. We also found a reduction of CD4 count level in the prednisolone arm. However, this reduction is not significant. As expected, patients' CD4 count levels significantly increases with increasing time, and patients who are on ART at each scheduled visit time have significantly higher CD4 count levels relative to those who were not on ART treatment at each scheduled visit. The near zero estimates of the prednisolone-time interaction effect suggest that there is no difference in prednisolone effect in both arms over time. The results also show that older patients are more likely to have lower CD4 values, hence CD4 count level significantly decreases

Table 5.2: Non-monotone pattern: Parameter estimates from *de jure* and *de facto* analyses (N = 857 patients)

	Prednisolone			**Time**			**Prednisolone x Time**			**ART**			**Prednisolone x ART**			**Age**		
Analysis[a]	Est	s.e.	p-value	Est	s.e.	p-value	Est.	s.e.	p-value	Est.	s.e.	p-value	Est.	s.e.	p-value	Est.	s.e.	p-value
MI	-0.61	0.42	0.150	0.51	0.079	<0.0001	-0.004	0.115	0.974	1.50	0.394	<0.0001	0.21	0.581	0.718	-2.51	0.667	< 0.0001
LMCF	-0.76	0.446	0.088	0.39	0.045	<0.0001	0.03	0.081	0.707	2.21	0.263	<0.0001	-0.04	0.406	0.924	-2.29	0.799	0.004
J2R (P-)	-0.36	0.486	0.455	0.47	0.047	<0.0001	0.04	0.076	0.561	2.03	0.258	<0.0001	-0.30	0.404	0.457	-2.25	0.779	0.004
J2R (P+)	-0.36	0.512	0.489	.45	0.052	<0.0001	0.03	0.080	0.737	2.06	0.321	<0.0001	-0.16	0.438	0.712	-2.25	0.795	0.001
CDR (P-)	-0.61	0.444	0.168	0.47	0.047	<0.0001	.05	0.083	0.577	2.04	0.259	<0.0001	-0.21	0.379	0.582	-2.27	0.783	0.0004
CDR (P+)	-0.59	0.471	0.216	0.46	0.051	<0.0001	0.02	0.083	0.849	1.93	0.288	<0.0001	-0.03	0.397	0.939	-2.24	0.790	0.005
CR (P-)	-0.64	0.438	0.143	0.47	0.047	<0.0001	0.04	0.075	0.644	2.04	0.258	<0.0001	-0.15	0.376	0.688	-2.29	0.782	0.003
CR (P+)	-0.66	0.451	0.142	0.46	0.050	<0.0001	0.02	0.079	0.792	1.91	0.288	<0.0001	-0.01	0.405	0.979	-2.22	0.792	0.001

[a]MI: multiple imputation; LMCF: last mean carried forward; J2R (P-): jump to placebo reference arm; J2R (P+): jump to active reference arm, CDR (P-): copy difference with placebo reference arm; CDR (P+): copy difference with active reference arm; CR (P-): copy reference with placebo reference arm; and CR (P+): copy reference with active reference arm

with increasing age. These results agree with the mean $\sqrt{\text{CD4}}$ count profiles plots in Figure 2.2 and Figure 2.4. This is because CD4 count in both the prednisolone and placebo arms increases at the same rate (we found no evidence that prednisolone influence the effect of ART) and CD4 count increases with increasing time where this increase, in both arms, is the same over time (no prednisolone-time effect).

5.6 Simulation study

Simulation studies are designed/conducted to achieve various aims. For instance White simulation studies focused on large or small sample bias and precision relative to other available methods (White and Thompson, 2005; White, 1997). Hughes et al. (2016) simulation studies investigated variance estimation and Morris et al. (2014) work investigated the robustness to mis-specification. Basically, some simulation studies are designed offer a proof-of-concept in order to show that a method is viable or fallible in some settings while others are designed to stress-test methods by identifying settings where the method may fail (Morris et al., 2019). These studies are very useful in statistical research. For instance, one may be faced with tow competing methods of analysis both of which are equally easy to implement. Even if the method chosen is unlikely to significantly influence the results of the analysis, it may be useful to have unrealistically extreme data-generating mechanisms to understand when and how each method fails (Morris et al., 2014). On the other hand, it may be useful to compare methods where some or all methods have been shown to work in principle but the methods under scrutiny were designed to address slightly different problems (Morris et al., 2019).

In this section we performed simulation experiments to evaluate the performance of the LMCF, J2R, CDR and CR methods under different rates of missing data. The hypothesis to address is whether these methods will be able to produce unbiased statistical inferences under different missingness rates. Under each missing data rate and missing data mechanism, we evaluate the bias of the usual MI method for imputation of missing data and likelihood based method (ML) against the LMCF, J2R, CDR and CR methods. The MI and ML methods are known to provide valid inference when missing values are missing at random (MAR) (Rubin, 1996, 1976).

The simulated datasets were generated using the R software. The R code for the simulation experiment is available from the first author upon request. The simulation experiment was performed according to the linear mixed effect model defined by

$$\begin{aligned} \mathrm{Y}_{ij} &= \beta_0 + \beta_1 \times \text{treatment}_i + \beta_2 \times \text{time}_j \\ &\quad + \beta_3 \times \text{treatment} \times \text{time}_{ij} + \mathrm{b}_i + \epsilon_{ij}. \end{aligned}$$

The initial values for $\beta_0, \beta_1, \beta_2$, and β_3 are 13, 0.75, 0.11, -0.19, 0.20 respectively. The initial value for standard deviation σ of the random effect b_i is 4.57 and that of the ϵ_i is 3.07. In generating these data sets, we assumed that (1) the measurement at the first time point (j=0) from the original data set is completely observed, (2) the data are MCAR or MAR mechanism, (3) the missing pattern is monotone, and (4) there are different dropout rates.

The probability that Y_{ij} is MCAR is $\Pr(R_{ij} = 1) = \vartheta_0$ and MAR is $\Pr(R_{ij} = 1) = \Phi(\vartheta_0 + \vartheta_1 Y_{i,j-1})$. We chose ϑ_0 so that the amount of missing data in the datasets are approximately 5%, 20%, and 50% for MCAR mechanism and ϑ_0 and ϑ_1 so that the missing data in the dataset are approximately 5%, 20%, and 50% for MAR mechanism. We considered the following two steps for generating the data sets. We called these steps, M-step and D-step. We generated the longitudinal measurements under the M-step and under the D-step, then generated data according to MAR and MCAR mechanisms.

M-step: We generated five-repeated measurements for each patient by a random number from a multivariate normal distribution. We used parameter estimates obtained from fitting a linear mixed effect model to the data. We repeated these processes 1000 times for 200 patients. Patients were randomly assigned to two treatment (treatment and placebo) arms in a ratio 1:1.

D-step: We generated missing data according to MCAR and MAR mechanisms. Missing data were generated through a logistic regression model. However, generating MCAR and MAR missing mechanisms involves two different assumptions for the dropout mechanism. For MAR, missing data were generated by dropping observations according to a logistic regression model relating the probability of dropout at particular time point with changes from baseline to previous time point. For MCAR, missing data were randomly generated by dropping observations according to a logistic regression model. Specific values for the logistic regression were chosen in order to yield the desire dropout rates in a given missing data mechanism. Under each of the missing data mechanisms, we generated overall dropout rates at 5%, 20%, 30, and then 50%. Thereafter, we performed analyses using ML, MI, LMCF, J2R, CDR and CR approaches and then assessed the performance of these methods in estimating treatment effect.

The results from the simulation study under MCAR and MAR mechanisms are shown in **Appendix** B. The MCAR results are shown in Table B.1 and the MAR results are presented in Table B.2. Under the MCAR mechanism, it can be observed that all the methods produced unbiased parameter estimates under the different missingness rates. The root mean square error (RMSE) estimates of treatment effect, produced by each method under the different missingness rates, are often higher compared with the time and treatment-time interaction effects. Most of the methods yielded unbiased estimates of treatment effect and this may imply that the process that generated the missing data is likely to be random.

The simulation results under the MAR mechanism revealed that each of the methods yielded un-

biased estimates for prednisolone effect under the missingness rates with less unbiased estimates for treatment effects when the missingness rates are 5%, 20% and 50%. All the methods yielded unbiased estimate of time effect under the different missingness rates. When missingness rate was assumed to be 50%, the LMCF and the CDR methods yielded less unbiased estimates of time effect. Each of the methods showed no bias for the treatment-time interaction slope when the missingness rates were assumed to be 5%, 10% and 30%, and bias for treatment-time interaction slope when the missingness rate was assumed to be 50% and 20%. However, ML and MI, yielded unbiased estimates for treatment-time interaction.

We have stated that the aim of the simulation studies is to evaluate the performance of LMCF, J2R, CDR and CR methods under different rates of missing data. Specifically, we investigate how the methods would perform under increasing rate (5%, 10%, 20%, or 50%) of missing data. Because the simulated datasets were generated using parameters informed by the IMPI trial, the simulation studies results show no significant treatment and treatment-time interaction effects and significant time effect.

We found that, even when the missingness rate was substantial as 50%, the methods still produced unbiased estimates with the exception of LMCF and CDR methods (which produced less unbiased estimates for time effect). This is expected for the LMCF method because it does not account for variability due to imputation of missing data and hence produces large number of identical imputed data as the amount of missing data increase. Since the CDR method assumes that after a patient deviates that patient's post-deviation mean increments copy those from the reference arm, this increments are likely to be identical with higher amount of missing data.

The J2R method appears to be robust to higher rate of missing values because after a patient stops taking treatment from the randomized arm, such patient's mean response distribution is now considered to be the same as that of the "reference" group of patients. By this assumption, the J2R methods is able to account for variability due imputation missing data (even with small amount information or substantial amount of missing data). This explained why the CR method also appears to be robust to substantial amount of missing data since a patient's whole distribution, both pre-deviation and post-deviation data, is the same as reference arm.

These findings suggest that the J2R and CR should be used when dealing with data with substantial amount of missing data. However, the CDR can also be used because a study with amount of missing data above 50% is too high. Thus, these methods (proposed by Carpenter et al. (2013)) can be used for handling the missing data in the IMPI clinical trial and other trials with similar settings. By definitions, the CR and J2R methods are recommended for use in the IMPI trial setting and other trials with similar settings because patients in either arm are expected to obtain equal benefit from ART.

5.7 Discussion

In this chapter, we conducted sensitivity analysis to investigate sensitivity of statistical inferences under MAR analysis (*de jure* option) to alternative plausible assumptions under NMAR (*de facto* option) using the PM-MI approach (Carpenter et al., 2013). In this thesis, we applied the PM-MI approach to incomplete longitudinal data. We also conducted simulation studies to evaluate the performance of the approach. The principles and methods considered quantify the robustness of inferences to departures from the primary analysis assumptions.

The PM-MI approach was implemented to the CD4 count data to investigate the effect of TB pericarditis treatment (prednisolone) on CD4 count changes over time. The study results show that inferences under the *de jure* (MAR primary analysis) assumption are robust to the inferences under the *de facto* (NMAR sensitivity analysis) assumptions. This finding gives an indication that the mechanism that generated the missing values in the CD4 count measurements from the IMPI trial is likely to be missing at random (MAR). The implications are that (1) the observed data are random sample from the population patients with TB pericarditis and (2) either the direct maximum likelihood (ML) approach or the multiple imputation approach, under the assumption that the data are MAR, can be used to produce valid inferences.

The investigation of sensitivity of statistical inferences to missing data is important, and use of such methods must be encouraged. This is because such sensitivity analysis provides additional information to readers of a clinical report to be able to interpret the results. This means that clinical reports should describe the primary and the sensitivity analyses to non-statisticians. This requires that assumptions about missing data are articulated in a transparent manner so that researchers and practicing clinicians can assess their validity under the study at hand (Carpenter et al., 2013). Carpenter et al. (2013) encourage the need for such sensitivity analysis stating that "assumptions need to be accessible, so that in the context of the trial at hand all stakeholders can understand whether or not they are plausible. Then, departures from these assumptions also need to be relevant in the context of the trial at hand, so that stakeholders can see if they require investigation." When data are missing, it is possible that readers of a clinical report may doubt the conclusions reached unless the conclusions are supported with sensitivity analysis.

Our study results from both the monotone (in Table 5.1) and the non-monotone (in Table 5.2) data showed that there is no significant prednisolone effect in all the analyses. The prednisolone-time interaction results show a very slight reduction in CD4 count level among the patients in the prednisolone arm, compared with the placebo arm over time. However, this reduction is not significant. As expected, there is a significant time effect, indicating that CD4 count level increases with increasing time. Patients who were on ART treatment, at each scheduled visit, are likely to

have significantly higher CD4 count levels compared with those who were not always on ART at each visit time. The results also show that older patients are more likely to have a lower CD4 count level. Also, there is no prednisolone-ART interaction effect in all the analyses. However, the prednisolone effects under the non-monotone analyses are negatives because the overall reduction in the CD4 count levels among patients in the prednisolone arm is more pronounced than that of the patients in the placebo arm (see Figure 2.2). On the contrary, the treatment effects under the non-monotone analyses are positives because the overall reduction in the CD4 count levels among patients in the prednisolone arm is less pronounced than that of the patients in the placebo arm (see Figure 2.4).

An alternative sensitivity analysis approach (known as re-weighting approach) (Carpenter et al., 2007), based on MI, is to assess sensitivity of statistical inferences under MAR primary analysis by weighting imputations under MAR mechanism to reflect a particular NMAR mechanism considered. Unlike the PM-MI approach (Carpenter et al., 2013), this approach avoids imputation under the NMAR mechanism. The ideas in the re-weighting approach were proposed by Carpenter et al. (2007). Rezvan et al. (2015) discussed the short falls of the re-weighting approach. These authors noted that this approach still suffers bias and that such bias should be recognized by users, and more appropriate methods should be developed. Smuk (2015) proposed a partial solution to improve the performance of the approach. Smuk (2015) could not give definitive guidelines as to when to apply the re-weighting method. This is because these methods are data-dependent. However, it was suggested that the method should be applied and compared to the re-weighting approach proposed by Carpenter et al. (2007).

The PM-MI sensitivity analysis approach Carpenter et al. (2013) was applied to measurements at the last visit (assuming that measurements are independent). So the linear regression model, that assumes that the measurements are independent, was used. In this thesis, we considered the PM-MI for incomplete longitudinal measurements, (measurement at all visits), assuming that such measurements are correlated using the linear mixed effects model for the correlated measurements. We also carried out simulation studies to evaluate the performance of the PM-MI (Iddrisu and Gumedze, 2019*a*). The simulation results revealed that the PM-MI approach is suitable for handling missing in the IMPI trial setting and other trials with similar settings and that the PM-MI approach is likely to produce bias estimates for some parameter estimate with increasing missingness rate.

Chapter 6

Sensitivity analysis for the generalized shared-parameter model for incomplete longitudinal data

6.1 Introduction

In this chapter, we considered alternative approaches for conducting sensitivity analyses to missing data. These sensitivity analyses approaches are within the shared-parameter model (SPM) framework introduced in Chapter 3. Unlike the selection model framework which assumes that the dropout probabilities depend on the response variable (where the measurement and the dropout models are joint/link response variable), the SPM framework assumes that the dropout probabilities depend on a set of common random effects (where the measurement and the dropout models are joint/link by these random effects).

In this chapter, we introduced the conventional SPM framework (Follmann and Wu, 1995*a*; Albert and Follmann, 2000) and the generalized SPM (GSPM) (Creemers et al., 2010). The GSPM is an extension of the SPM discussed in Section 3.1.3. We discuss the specification of the generalized shared-parameter model (GSPM) and also explain how the shared random effects make it possible to perform various forms of sensitivity analyses (Creemers et al., 2011). We proposed and conducted sensitivity analyses for the GSPM framework using the global influence approach. The global approach with the GSPM allows one to investigate the effect of potentially influential subjects on the parameter estimates, dropout mechanism, and model key conclusions (Cook, 1986; Molenberghs and Kenward, 2007). In Section 6.3 and Section 6.4, these approaches were implemented to the CD4 count data introduced in Chapter 2, Section 2.6 and the results were discussed. We also conducted simulation studies to evaluate the performance of the GSPM.

The shared-parameter model is a joint model for non-ignorable missing data (NMAR) mechanism (Rubin, 1976; Wu and Bailey, 1988; Wu and Carroll, 1988). As we have stated in our previous discussions, it is known that valid inference can be obtained when the process by which missing data are generated is assumed to be missing at random (MAR). However, such an assumption would produce biased inference if the missing data generating process is not missing at random

(NMAR). This later assumption requires a joint model for the measurement and the missingness process in order to produce valid inference. There is, therefore, the need to perform sensitivity analysis to assess the sensitivity of the inference from the MAR primary analysis to alternative plausible assumptions under the NMAR sensitivity analyses.

Various authors have proposed a joint model for the measurement and the dropout processes. The shared-parameter modeling framework (Wu and Bailey, 1988; Wu and Carroll, 1988; Tsonaka et al., 2009; Creemers et al., 2010) is one approach that offers and appropriate framework for joint modeling of the measurement and the missingness processes. One important feature of this approach is that it assumes that the measurement and the missingness models share a set of random effects (Albert et al., 2002; Roy, 2003; Creemers et al., 2010; Mallinckrodt, Lin and Molenberghs, 2013). These random effects account for the association between the measurement and the missingness models. These random effects also account for the correlation between the repeated measurements.

In Wu and Carroll (1988) study of repeated measurements of lung function, they proposed the SPM in which the measurement process was modeled using the linear mixed effect model (LMM) (Laird and Ware, 1982) with a random intercept and slope. They assumed a probit model for the censoring process. The LMM was linked with the censoring process model by including a patient's random slope as a covariate in the probit model for the censoring process. In this way, when the probit regression coefficient for the random slope is non-zero, it gives an indication that there is dependence between the measurement and the missingness processes. This implies that any method of analysis that does not account for this dependence may produce biased inferences. Albert and Follmann in (Fitzmaurice et al., 2008, Ch. 19) showed that using such naive analysis will tend to weight the complete cases more heavily than they should have been weighted.

The random effects in the SPM reflect patients' deviation from the mean estimates of the fixed effects. Each patient draws a random slope from a Gaussian distribution. This slope governs both the patient's expected rate of decline in the response and the probability of dropping out (Molenberghs and Fitzmaurice in (Fitzmaurice et al., 2008, Ch.19)). Albert et al. (2002) also pointed out that the shared-parameter models make assumptions which can only in part be justified using the data. For instance, one can determined whether, for the observed data, change in response follows a linear function of time or a quadratic function of the log time, using standard diagnostics tool such as examining the residual correlation. On the contrary, changes in the response for the unobserved data, which are imposed by the SPM, cannot be justified. This can only be verified by performing sensitivity analysis. Creemers et al. (2010) proposed the generalized shared-parameter model (GSPM) and Creemers et al. (2011) introduced sensitivity analysis tools within the GSPM framework.

The choice of measurement model depends on the type longitudinal data being analyzed. For

normally distributed longitudinal data, Wu and Bailey (1989) assumed the linear mixed model (Laird and Ware, 1982). The responses are assumed to follow the normal distribution with a defined mean and variance-covariance structure. For a discrete or dichotomous longitudinal data (Follmann and Wu, 1995*b*; Pulkstenis et al., 1998; Wu and Follmann, 1999), the response model can be formulated as a generalized linear mixed effect model (Follmann and Wu, 1995*a*; Thomas et al., 1998; Albert and Follmann, 2000; Alfò and Aitkin, 2000). Albert and Follmann (2000) proposed methodology for longitudinal count data.

The type of model for the missingness process depends on the type of missing data being considered. For instance, when the data are a discrete time to dropout, then a geometric distribution can be used to model the missingness process (Mori et al., 1994). Several authors have proposed the SPM for the case where dropout is a continuous event time (Schluchter, 1992; De Gruttola and Tu, 1994; Tsiatis et al., 1995; Tsiatis and Davidian, 2001; Vonesh et al., 2002).

6.2 Shared-parameter model framework

The focus of this section is the SPM framework. We will consider the generalized shared-parameter model (GSPM) (Creemers et al., 2010, 2011). The GSPM is the most general SPM (random-effects model) that allows different components of the SPM to have different set of random effects. That is, some of the random effects are shared between a pair of components and others are restricted to a single component.

6.2.1 Shared-parameter model under the IMPI trial

In our analysis we fitted a linear mixed for the CD4 count with random patient effects for intercept and slope. The random slope governs the patient's expected rate of change in the outcome and the probability of dropping out (see Molenberghs and Fitzmaurice in Fitzmaurice et al. (2008), Ch. 19). Hence, it is important consider a model for the dropout probability such that the probability of dropout depends on the random slope rather than the response (as seen in DK selection model Diggle and Kenward (1994)). The SPM framework provides a tool for joint modeling of the measurement and the dropout processes. It is assumed that the measurement process model and the dropout process model share a set of random effects. These random effects account for the association between the measurement model and the missingness model and also account for the correlation between the repeated measurements. It allows us to make different assumptions regarding the association between the measurement model and the dropout model (via the random effects), and then compare the results to assess the robustness of the statistical inferences to these assumptions.

6.2.2 Notation and concepts

Let N be the number of subjects in the study with n_i-dimensional vector of the response $\mathbf{Y}_i = (Y_{i1}, Y_{i2} \ldots, Y_{in})$ measured for each subject $i, i = 1 \ldots N$ at occasion $j, j = 1 \ldots n_i$, where n_i is the time at which the ith subject dropped out. Also let $\mathbf{X}_i$ be an $N \times p$ design matrix of covariates associated with the response and indicator with $\mathrm{R}_{ij} = 1$ if y_{ij} is observed and 0 otherwise and $\mathbf{b}_i$ be vector of random effects. The density function of the data is $\Pr(\mathbf{Y}_i, \mathbf{R}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi})$, where the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ describe the measurement and the missingness processes respectively. The conventional shared-parameter model (SPM) (Fitzmaurice et al., 2008, Ch. 19), with covariates information suppressed, is defined as

$$\Pr(\mathbf{Y}_i, \mathbf{R}_i, \mathbf{b}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}, \mathbf{G}_i(\boldsymbol{\rho})) = \Pr(\mathbf{Y}_i \mid \mathbf{b}_i, \boldsymbol{\theta}) \times \Pr(\mathbf{R}_i \mid \mathbf{b}_i, \boldsymbol{\psi}) \times \Pr(\mathbf{b}_i \mid \mathbf{G}_i(\boldsymbol{\rho})), \tag{6.1}$$

where $\mathbf{G}_i(\boldsymbol{\rho})$ is covariance matrices of the random effects $\mathbf{b}_i$. Also see Tsonaka et al. (2009); Ibrahim and Molenberghs (2009); Follmann and Wu (1995a) for details on the SPM formulations and parameters estimation.

Creemers and colleagues (Creemers et al., 2010) expanded the $\mathbf{b}_i$ to be a set of latent structures which forms the basis for the general SPM (GSPM). Let $\mathbf{Y}_i^o$ and $\mathbf{Y}_i^m$ denote observed and unobserved components of $\mathbf{Y}_i$ respectively. The GSPM (Creemers et al., 2010) assumes a set of random effects vectors $\mathbf{b}_i = (\mathbf{g}_i, \mathbf{h}_i, \mathbf{j}_i, \mathbf{k}_i, \mathbf{l}_i, \mathbf{m}_i, \mathbf{q}_i)$ and the full density function is defined as

$$\begin{aligned}\Pr(\mathbf{Y}_i, \mathbf{R}_i \mid \mathbf{b}_i, \boldsymbol{\theta}, \boldsymbol{\psi}) = \Pr(\mathbf{Y}_i^o \mid \mathbf{g}_i, \mathbf{h}_i, \mathbf{j}_i, \mathbf{l}_i, \boldsymbol{\theta}) \times \Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{g}_i, \mathbf{h}_i, \mathbf{k}_i, \mathbf{m}_i, \boldsymbol{\theta}) \\ \times \Pr(\mathbf{R}_i \mid \mathbf{g}_i, \mathbf{j}_i, \mathbf{k}_i, \mathbf{q}_i, \boldsymbol{\psi}),\end{aligned} \tag{6.2}$$

where $\mathbf{g}_i, \mathbf{h}_i, \mathbf{j}_i, \mathbf{k}_i, \mathbf{l}_i, \mathbf{m}_i, \mathbf{q}_i$ are $N \times q$ dimensional independent random-effects vectors. These random effects are normally distributed with mean zero and unknown variance. This is the most general SPM in which the random effect $\mathbf{g}_i$ is shared among the three components (eg., $\mathbf{Y}_i^o$, $\mathbf{Y}_i^m$, and $\mathbf{R}_i$ models) in the GSPM (6.2). The random effects, $\mathbf{h}_i, \mathbf{j}_i$ and $\mathbf{k}_i$, are shared between a pair of components and then, $\mathbf{l}_i, \mathbf{m}_i$ and $\mathbf{q}_i$ are restricted to a single component. The random effect $\mathbf{m}_i$ is not identifiable because it only describes the missing data mechanism. Also, the random effect $\mathbf{k}_i$ is never identifiable because it is aliased with the random effect $\mathbf{q}_i$ of which one is used twice and the other in a single component only (Creemers et al., 2011, 2010). However, occurrence of $\mathbf{k}_i$ in the middle component does not separate it from $\mathbf{q}_i$ because the middle component is not identifiable. Also, the $\mathbf{g}_i$ and $\mathbf{j}_i$ are not separable. This means that the random effects $\mathbf{k}_i, \mathbf{q}_i, \mathbf{j}_i$ and $\mathbf{g}_i$ can be used as a means of conducting sensitivity analysis (Creemers et al., 2011). A special case of the GSPM is the conventional SPM (Wu and Bailey, 1988; Wu and Carroll, 1988; Wu and Bailey, 1989; Little, 1995; Follmann and Wu, 1995a) is defined as

$$\begin{aligned}\Pr(\mathbf{Y}_i, \mathbf{R}_i \mid \mathbf{g}_i, \boldsymbol{\theta}, \boldsymbol{\psi}) &= \Pr(\mathbf{Y}_i \mid \mathbf{g}_i, \boldsymbol{\theta}) \times \Pr(\mathbf{R}_i \mid \mathbf{g}_i, \boldsymbol{\psi}) \\ &= \Pr(\mathbf{Y}_i^o \mid \mathbf{g}_i, \boldsymbol{\theta}) \times \Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{g}_i, \boldsymbol{\theta}) \times \Pr(\mathbf{R}_i \mid \mathbf{g}_i, \boldsymbol{\psi}),\end{aligned} \tag{6.3}$$

Table 6.1: Sub-models of the general shared-parameter model

GSPM	Model	MAR counterpart
M_1	$\Pr(\mathbf{Y}_i^o \mid \mathbf{g}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{g}_i)\Pr(\mathbf{R}_i \mid \mathbf{g}_i)$	$\Pr(\mathbf{Y}_i^o)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i)$
M_2	$\Pr(\mathbf{Y}_i^o \mid \mathbf{g}_i, \mathbf{h}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{g}_i, \mathbf{h}_i)\Pr(\mathbf{R}_i \mid \mathbf{g}_i)$	$\Pr(\mathbf{Y}_i^o)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i)$
M_3	$\Pr(\mathbf{Y}_i^o \mid \mathbf{g}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{g}_i, \mathbf{k}_i)\Pr(\mathbf{R}_i \mid \mathbf{g}_i, \mathbf{k}_i)$	$\Pr(\mathbf{Y}_i^o)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i)$
M_4	$\Pr(\mathbf{Y}_i^o \mid \mathbf{g}_i, \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{g}_i, \mathbf{j}_i)\Pr(\mathbf{R}_i \mid \mathbf{g}_i, \mathbf{j}_i)$	$\Pr(\mathbf{Y}_i^o \mid \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i \mid \mathbf{j}_i)$
M_5	$\Pr(\mathbf{Y}_i^o \mid \mathbf{h}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{h}_i, \mathbf{k}_i)\Pr(\mathbf{R}_i \mid \mathbf{k}_i)$	$\Pr(\mathbf{Y}_i^o)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i)$
M_6	$\Pr(\mathbf{Y}_i^o \mid \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{k}_i)\Pr(\mathbf{R}_i \mid \mathbf{j}_i, \mathbf{k}_i)$	$\Pr(\mathbf{Y}_i^o \mid \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i \mid \mathbf{j}_i)$
M_7	$\Pr(\mathbf{Y}_i^o \mid \mathbf{h}_i, \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{h}_i)\Pr(\mathbf{R}_i \mid \mathbf{j}_i)$	$\Pr(\mathbf{Y}_i^o \mid \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i \mid \mathbf{j}_i)$
M_8	$\Pr(\mathbf{Y}_i^o \mid \mathbf{h}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{h}_i)\Pr(\mathbf{R}_i)$	$\Pr(\mathbf{Y}_i^o)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i)$
M_9	$\Pr(\mathbf{Y}_i^o \mid \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i \mid \mathbf{j}_i)$	$\Pr(\mathbf{Y}_i^o \mid \mathbf{j}_i)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i \mid \mathbf{j}_i)$
M_{10}	$\Pr(\mathbf{Y}_i^o)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{k}_i)\Pr(\mathbf{R}_i \mid \mathbf{k}_i)$	$\Pr(\mathbf{Y}_i^o)\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o)\Pr(\mathbf{R}_i)$

where only one vector of random effects $\mathbf{g}_i$ is assumed, conditional upon which the measurement and the dropout processes are independent. Various characterizations of the GSMP in the MCAR, MAR and NMAR frameworks can be found in (Creemers et al., 2011, 2010).

6.3 Sub-models of the GSPM model

In this section, we will analyze the longitudinal profile of CD4 count measurements using various sub-models of the GSPM model (6.2). These sub-models represent different assumptions about the dropout process via the random effects and thus allow us to compare inferences across the various dropout assumptions. The sub-models are obtained by removing portions of the random effects structure (Creemers et al., 2011). Table 6.1 presents the different forms of sub-models together with their MAR counterparts (Creemers et al., 2010).

Given such collection of sub-models, one particular model can be selected because of its interpretation, or one or a few models can be excluded for this same reason. In particular, M_{10} assumes that the observed measurements are independent. However, such assumption is not plausible for the CD4 count data and thus is excluded from the analyses. In addition, models can be considered jointly for the purpose of sensitivity analysis which is the focus of this paper.

Although we will discuss the results obtained from fitting M_1-M_9, we present only results from the M_1 and M_6 for illustration and sensitivity analysis.

The linear mixed effects model (LMM) (Laird and Ware, 1982) is assumed for the measurement process and is given by

$$\begin{cases} \mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \\ \mathbf{b}_i \sim N\left(\mathbf{0}, \mathbf{G}_i(\boldsymbol{\rho})\right), \\ \boldsymbol{\epsilon}_i \sim N\left(\mathbf{0}, \mathbf{R}_i(\boldsymbol{\sigma})\right), \\ \mathbf{b}_i \perp\!\!\!\perp \boldsymbol{\epsilon}_i, \end{cases} \tag{6.4}$$

where $\mathbf{b}_i$ is an q-dimensional vector of random effects, $\mathbf{Z}_i$ and $\mathbf{X}_i$ are $N \times p$ and $N \times p$ dimensional matrices of known covariates, $\boldsymbol{\beta}$ is a p-dimensional vector containing the fixed effects, $\boldsymbol{\epsilon}_i$ is an N-dimensional vector of residual components, $\mathbf{G}_i(\boldsymbol{\rho})$ and $\mathbf{R}_i(\boldsymbol{\sigma})$ are $q \times q$ and $n_i \times n_i$ covariance matrices respectively and $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ are $c \times 1$ and $s \times 1$ (with $s \leq n_i(n_i+1)/2$) vectors of unknown variance parameters corresponding to $\boldsymbol{\epsilon}_i$ and $\mathbf{b}_i$ respectively.

Let T_i, t_j, a_{ij}, and c_i denote prednisolone, month, anti-retroviral therapy-ART, and age variables respectively. These variables are included in the mean structure together with prednisolone $\times$ month and prednisolone $\times$ interactions. We assume the unstructured variance-covariance matrix for $\mathbf{R}_i$. An overview for mean and variance-covariance matrix of the random effects, in the case of the M_{1-9}, are presented in Table 6.2. The $\rho_g^2, \rho_k^2, \rho_h^2$, and ρ_j^2 are the variance parameters associated with the random effects $\mathbf{g}_i, \mathbf{k}_i, \mathbf{h}_j$, and $\mathbf{j}_i$ respectively.

A logistic regression model is assumed for the missingness process with the same parametric structure as for the mean structure. The various forms of the logistic regression model are presented in Table 6.3. The parameter estimates, γ_g, γ_k, and γ_j are scale factors for the shared random effects in the missingness model. The purpose of these scale factors is to avoid forced equality of the variance in the measurement and dropout model (Creemers et al., 2011). As a result, we have $\gamma_g g_i \sim N\left(0, \gamma_g^2\rho_g^2\right), \gamma_k k_i \sim N\left(0, \gamma_k^2\rho_k^2\right)$, and $\gamma_j j_i \sim N\left(0, \gamma_j^2\rho_j^2\right)$.

We fitted the models in this paper using the SAS procedure NLMIXED (Creemers et al., 2011) attached in **Appendix** C. The empirical Bayes estimates were used to obtain predictions for the incomplete profiles.

Although we have presented parameter estimates for the models (M_1 and M_6) in Table 6.4, the discussion is based on the results obtained from all the models (M_1-M_9). The parameter estimates for the mean structure remain unchanged in all the models. We observed that the prednisolone and prednisolone $\times$ ART interaction effects are not significant and CD4 count level increases significantly over time. The negative estimates for the sensitivity parameters implies that, with an increase of CD4 count level, patients are more likely to leave the study. This means that patients are more likely to leave the study as they become healthier or fully cured. However, this is not entirely correct in the case of HIV patients, since patients can only become healthier but not cured, and the CD4 count level is more likely to decline after the patient stops taking ART.

Table 6.2: Models for CD4 count measurements data[a]

Model	$\mathrm{E}[\mathbf{Y}^o \mid \mathbf{b}_i]$	$\mathrm{E}[\mathbf{Y}^m \mid \mathbf{Y}^o, \mathbf{b}_i]$	**BD**[b]
M_1	$\beta_0 + g_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + g_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	ρ_g^2
M_2	$\beta_0 + g_i + h_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + g_i + h_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\begin{pmatrix} \rho_g^2 & 0 \\ 0 & \rho_h^2 \end{pmatrix}$
M_3	$\beta_0 + g_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + g_i + k_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\begin{pmatrix} \rho_g^2 & 0 \\ 0 & \rho_k^2 \end{pmatrix}$
M_4	$\beta_0 + g_i + j_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + g_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\begin{pmatrix} \rho_g^2 & 0 \\ 0 & \rho_j^2 \end{pmatrix}$
M_5	$\beta_0 + h_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + h_i + k_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\begin{pmatrix} \rho_h^2 & 0 \\ 0 & \rho_k^2 \end{pmatrix}$
M_6	$\beta_0 + j_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + k_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\begin{pmatrix} \rho_j^2 & 0 \\ 0 & \rho_k^2 \end{pmatrix}$
M_7	$\beta_0 + h_i + j_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + h_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\begin{pmatrix} \rho_j^2 & 0 \\ 0 & \rho_h^2 \end{pmatrix}$
M_8	$\beta_0 + h_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + h_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	ρ_h^2
M_9	$\beta_0 + j_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	$\beta_0 + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$	ρ_j^2

[a] *Mean structures for the observed responses given the random effects, mean structures for the missing responses given the observed responses, and the random effects and the variance-covariance matrix for the random effects for the various Models. T_i is the treatment indicator for subject i and t_i is the time at which the j^{th} measurement is taken*

[b] BD: block diagonal matrix.

Table 6.3: Models for missingness mechanisms

Model	$\text{logit}[\Pr(R_{ij}=1 \mid R_{i,j-1}=0, \mathbf{b}_i, T_i, t_j, \gamma)]$
M_1	$\gamma_0+\gamma_g g_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_2	$\gamma_0+\gamma_g g_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_3	$\gamma_0+\gamma_g g_i+\gamma_k k_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_4	$\gamma_0+\gamma_g g_i+\gamma_j j_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_5	$\gamma_0+\gamma_k k_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_6	$\gamma_0+\gamma_j j_i+\gamma_k k_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_7	$\gamma_0+\gamma_j j_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_8	$\gamma_0+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$
M_9	$\gamma_0+\gamma_j j_i+\gamma_1 T_i+\gamma_2 t_j+\gamma_3 T_i t_j+\gamma_4 a_{ij}+\gamma_5 a_{ij}T_i+\gamma_6 c_i$

The scale parameters γ_k and γ_j are never identifiable and hence were fixed to 1 in these analyses. However, the unidentifiability of γ_k and γ_j provides a tool to perform sensitivity analysis (Creemers et al., 2011) in Section 6.4.

6.4 Sensitivity analysis approaches

In this section, we discuss and conduct sensitivity under the general shared-parameter model (GSPM) (Creemers et al., 2011, 2010). We will conduct sensitivity analysis and present the results of the models M_1 and M_6. For the rational of why only the models M_1 and M_6, we note that given the collection of sub-models M_1-M_{10}, one may decide to use a particular sub-model because of its interpretation. On the other hand, one or a few sub-models cannot be considered due to their interpretations or specifications. For sensitivity analysis purpose, models can be considered jointly for the purpose of sensitivity analysis which is the focus of this paper. Since the sub-models M_1-M_9 are all plausible under the IMPI trial setting, we conducted sensitivity using the sub-models M_1-M_9. Since the statistical inferences under these sub-models are comparable, the results under the sub-model M_1 (which is the conventional shared-parameter model) and M_6 (in Table 6.4) with relatively large variance for the random effect k_i were selected for discussion and further statistical assessment. We conducted simulation studies to evaluate the performance of the GSPM using the M_6. We propose to assess the impact of potentially influential subjects, on statistical inferences, using the global influence approach.

Table 6.4: CD4 count data: parameter estimates (Est) and standard errors (SE) for the SPM fits

Effects	Parameter	M_1	M_6
Mean structure		Est(SE)	Est(SE)
Measurement			
prednisolone	β_1	0.21 (0.86)	0.21 (0.856)
Month	β_2	0.26 (0.11)	0.28 (0.11)
prednisolone × Month	β_3	-0.03 (0.15)	-0.02 (0.15)
ART	β_4	3.03 (0.58)	2.98 (0.58)
prednisolone × ART	β_5	-0.23 (0.80)	-0.23 (0.80)
Age	β_6	-3.21 (1.55)	-3.12 (1.55)
Dropout			
prednisolone	γ_1	0.45 (0.50)	0.80 (1.67)
Month	γ_2	0.67 (0.12)	1.82 (0.47)
prednisolone × Month	γ_3	-0.15 (0.14)	-0.34 (0.40)
ART	γ_4	-2.49 (0.54)	-5.78 (1.87)
prednisolone × ART	γ_5	0.51 (0.68)	1.91 (2.02)
Age	γ_6	0.08 (0.51)	0.04 (2.20)
Variance-covariance structure			
Measurement			
Resid.var.	σ^2	8.55 (0.57)	8.51 (0.57)
Rand.int.var.	ρ_g^2	20.34 (2.75)	
	ρ_h^2		
	ρ_k^2		27.37 (13.00)
	ρ_j^2		20.31 (2.74)
Dropout			
Scale factor	γ_g	-0.10 (0.04)	
	γ_k		$\gamma_k = 1$
	γ_j		-0.18 (0.14)
Rand.int.var.	$\gamma_g^2\rho_g^2$	0.19 (0.15)	0.206 (14.20)
	$\gamma_j^2\rho_j^2$		0.68 (1.05)

6.4.1 Sensitivity analysis base on γ_k

In this section, we assess sensitivity of inferences across different values of γ_k. Although the sensitivity analysis conducted in this section is implemented to the M_6, the approach can be applied to M_1- M_9. The model M_6 includes the random effects $\mathbf{j}_i$ and $\mathbf{k}_i$ and since the second term $\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{g}_i, \mathbf{h}_i, \mathbf{k}_i, \mathbf{m}_i, \boldsymbol{\theta})$ in the model (6.2) is never observed, the sensitivity parameter γ_k is never identified. Hence, the model can only be fitted by fixing the sensitivity parameter, γ_k (see the results under Table 6.4).

For sensitivity analysis, the principle is to allow γ_k to take different values (eg., $\gamma_k = k = [-10, 10]$) and compare inferences across the different values of γ_k. For a given value of γ_k, linear mixed effects model will be fitted and then, based on this model, the missing responses ($\mathbf{Y}_i^m$) will be imputed K times (Creemers et al., 2011). This will be done using the conventional multiple imputation approach Little and Rubin (1987); Rubin (1996); Molenberghs and Kenward (2007). This means that, for each value of γ_k, K data sets are imputed and then a model fitted to each of the K data sets yielding K statistics. Thereafter, these K statistics are combined, using Rubin's rule (Rubin, 1976, 1996), to a single set of inferences. Inferences are then compared across the different γ_k values to assess sensitivity of the results (Creemers et al., 2011).

6.4.2 Step by step implementation of the methodology

The implementation steps outlined in this section is applicable to all models with at least one unidentifiable sensitivity parameter (Creemers et al., 2011). Here, we applied the algorithm to the model M_6. We specified a range for the values of this sensitivity parameter γ_k, from -10 to 10, with step length 1. When graphically exploring the results, the user can decide to give more or less weight to certain regions of the sensitivity parameters, in accordance with prior belief or other scientific opinions held (Creemers et al., 2011).

The following steps were adopted for every fixed value in the range of the sensitivity parameter:

STEP 1: Fit the model for $\mathbf{Y}_i^o$ using the fixed value of γ_k, for $\gamma_k = [-10, 10]$.

STEP 2: Calculate the conditional distribution $\Pr(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, k_i, \boldsymbol{\theta})$.

STEP 3: Using the empirical Bayes estimates, in STEP 2, simulate K values for $\mathbf{Y}_i^m$, conditional on $\mathbf{Y}_i^o$ and the random effect k_i. This give rise to K different complete data sets.

STEP 4: Fit the linear mixed effect model to each of the K completed data sets.

STEP 5: Combine the results from these K data sets to form a single set of inferences.

The STEP 3 can be seen as multiple imputation under MNAR and thus enables the use of conventional multiple imputation to perform STEP 5. Because $\gamma_k = 0$ corresponds to MAR, repeating this procedure for the whole grid results in a sensitivity analysis around MAR in the GSPM framework. If results are very sensitive to the value of γ_k, one should be very careful interpreting them (Creemers et al., 2010).

We have used forest plots to visually display parameter estimates and their corresponding confidence intervals (CI) for different values of γ_k. Based on these plots, one will be able see if inference, on a given parameter estimate changes for different values of γ_k. Although we have displayed forest plots for only prednisolone in Figure 6.1 and ART in Figure 6.2, we will discuss the results of the forest plots for the other parameter estimates in the model.

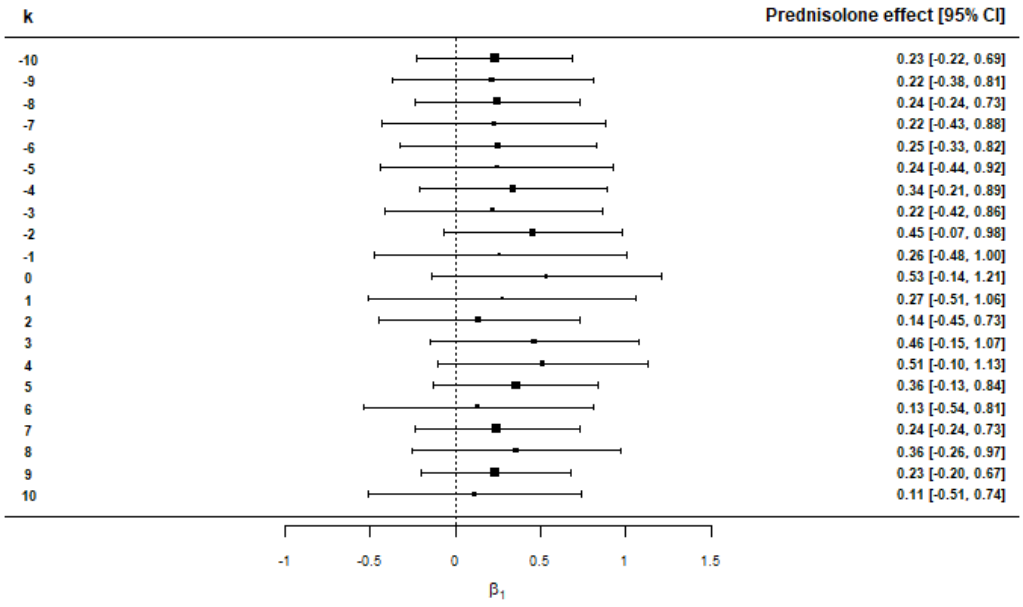

Figure 6.1: Forest plot of prednisolone effect β_1 and 95% confidence intervals for different values of the scale parameter ($k : \gamma_k$).

It can be observed that the prednisolone effect β_1 is not statistically significant, whereas the ART effect β_4 is statistical significant. The forest plots also showed that the prednisolone $\times$ time interaction effect β_3 and prednisolone $\times$ ART interaction effect β_5 are not significant whereas the time effect β_1 and age effect β_6 are statistically significant. The variance parameters σ^2 and ρ_g^2 are significant. The results showed that the parameter estimates remained unchanged for the different values γ_k. We observed that, over the entire range of values for γ_k, the CI for each of the parameter estimates remain consistent with no strong difference in evolution. Although the value of γ_k changes across models, the results and conclusions are in line with the results under the model M_6. The

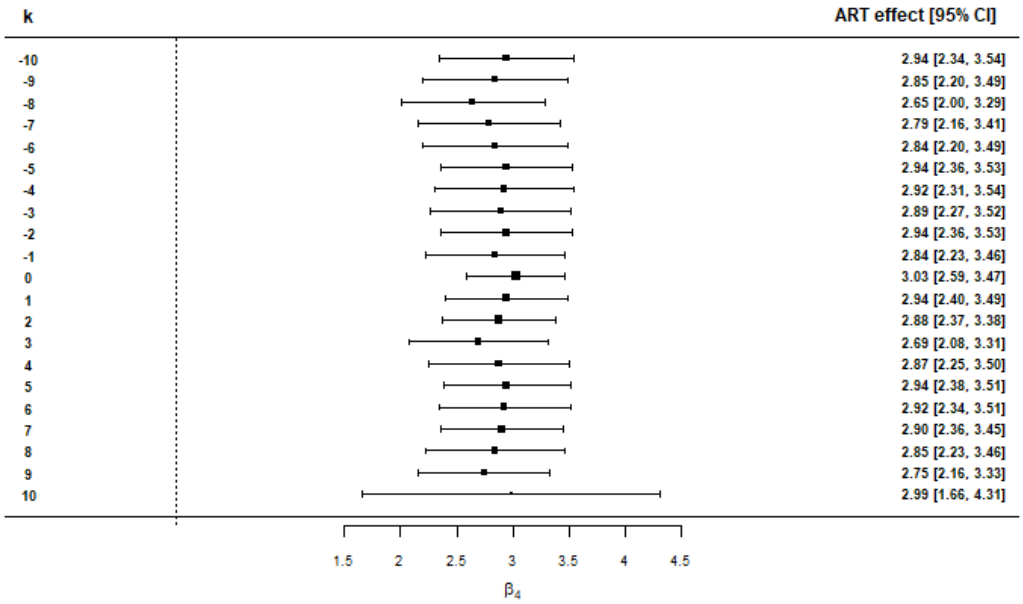

Figure 6.2: Forest plot of ART effect β_4 and 95% confidence intervals for different values of the scale parameter $(k : \gamma_k)$.

parameter γ_k is the main quantity which drives the sensitivity analysis and hence gives us some level of confidence in the results obtained and conclusions reached.

6.4.3 Global influence analysis

In this section we propose to assess the impact of potentially influential subjects, on the statistical inferences, using the global influence approach. This approach can be used within any of the models in Tables 6.1 and 6.2. However, we considered the M_1 for illustration purpose.

In the model M_1, the random effects g_i are shared by all the three components of the models for $\mathbf{Y}_i^o$, $\mathbf{Y}_i^m$, and $\mathbf{R}_i$. It is assumed that the measurement and missingness models, share common random effects (Gao, 2004; Wu and Bailey, 1988, 1989). These random effects account for the association between the measurement and the missingness models and the correlation between the repeated measurements. The model M_1 can then be written as

$$[\mathbf{Y}_i^o \mid \mathbf{b}_i] = \beta_0 + \mathbf{Z}g_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i. \tag{6.5}$$

The probability of dropout model can be written as

$$\begin{aligned}\text{logit}\left[\Pr\left(R_{ij} = 1 \mid R_{i,j-1} = 0, \mathbf{b}_i, T_i, t_j, \gamma\right)\right] = \gamma_0 + \mathbf{Z}g_i + \gamma_1 T_i + \gamma_2 t_j + \gamma_3 T_i t_j \\ + \gamma_4 a_{ij} + \gamma_5 a_{ij} T_i + \gamma_6 c_i.\end{aligned} \tag{6.6}$$

Firstly, we investigate the dependence between the measurement model and the dropout model by using the variance $\mathbf{G}(\boldsymbol{\rho})$ or the standard deviation of the random effect $\mathbf{g}_i$. That is, when the

$\boldsymbol{\rho} \neq \mathbf{0}$ (NMAR), the measurement model depends on the dropout model via $\mathbf{g}_i$ and hence parameter estimates are expected to change considerably. However, when $\boldsymbol{\rho} = \mathbf{0}$ (MAR), the measurement and the dropout models no longer share common random effects and parameter estimates are expected to remain unchanged under varying missingness assumptions.

For sensitivity analysis purpose, we consider three scenarios under the model M_1 (Tables 6.1 and 6.2). The Scenario 1 assumes that the measurement model and the dropout model do not share a common random effect (MAR) and Scenario 2 assumes that the measurement model and the dropout model share a common random intercept (NMAR). Scenario 3 assumes that the measurement model and the dropout model share common random intercept and time slope terms (another form of NMAR mechanism). The scenarios 1-3 formulations allow us to specify the shared-parameter model in such a way that the some of the parameters in both the dropout model and the measurement model share common random effects. These formulations allowed us to assess the effect of the shared-random effects on the dropout mechanism as well as statistical inferences to be drawn.

We perform analyses under these scenarios and then compare the results. The results from these analyses are presented in Table 6.5. The parameter estimates remained unchanged under the three models (Scenario 1, Scenario 2, and Scenario 3) and statistical inferences about the parameter estimates are in line with that of the statistical inferences in Table 6.4. Also, the log-likelihood value for the Scenario 1 (which is equivalent to MAR assumption) in both tables is highest. This suggests that models that assume that the CD4 count values are MAR are more likely to provide best fits for the data.

Now, we introduce the global influence approach and then implement this approach to study the impact of one or more subjects, with large influence, that may overturn the statistical inferences (Cook, 1986; Molenberghs and Kenward, 2007). This sensitivity analysis tool starts from case-deletion and is based on the difference in log-likelihood between the models fitted to the entire dataset (Molenberghs and Kenward, 2007, pp. 386-387). The log-likelihood function for the measurement model and the dropout model can be written as

$$\ell(\boldsymbol{\varphi}) = \sum_{i=1}^{N} \ell_{(i)}(\boldsymbol{\varphi}),$$

where $\ell_{(i)}(\boldsymbol{\varphi})$ is the contribution of the ith subject to the log-likelihood and where $\boldsymbol{\varphi}\prime = (\boldsymbol{\beta}\prime, \boldsymbol{\alpha}\prime, \boldsymbol{\delta}\prime)$ is the vector of the parameters in the measurement model and the dropout model and $\ell_{(-i)}(\boldsymbol{\varphi})$ is the log-likelihood function after the ith subject has been removed. Cook's Cook (1986) distances (CD) are based on measuring the discrepancy between the maximum likelihood $\ell(\boldsymbol{\varphi})$ and $\ell_{(-i)}(\boldsymbol{\varphi})$ or subset of the estimated parameter vectors $\hat{\boldsymbol{\varphi}}$ and $\hat{\boldsymbol{\varphi}}_{(-i)}$, where $\hat{\boldsymbol{\varphi}}_{(-i)}$ is the maximum likelihood estimate after the ith subject has been removed. The CD is defined as (Molenberghs and Kenward,

Table 6.5: Parameter estimates, standard errors, and p-values for the Scenarios 1-3.

Model	Scenario 1			Scenario 2			Scenario 3		
Parameter	Est.	s.e	p-value	Est.	s.e	p-value	Est.	s.e	p-value
Measurement model									
Time	0.43	0.12	0.001	0.51	0.14	0.001	0.48	0.14	0.001
Prednisolone	0.27	0.87	0.760	0.26	0.87	0.765	0.22	0.87	0.799
Prednisolone × Time	-0.09	0.16	0.567	-0.07	0.17	0.675	-0.03	0.19	0.876
ART	2.66	0.56	0.001	2.61	0.56	0.001	2.61	0.56	0.001
Prednisolone × ART	-0.29	0.76	0.701	-0.28	0.76	0.712	-0.30	0.76	0.690
Age	-2.94	1.55	0.060	-2.95	1.55	0.059	-2.94	1.53	0.056
Dropout model [a]									
γ_0	-	-	-	0.02	0.06	0.684	0.10	0.08	0.177
ρ_{γ_0}	-	-	-	1.50	1.06	0.158	0.91	1.37	0.508
γ_1	-	-	-	-	-	-	0.04	0.08	0.588
ρ_{γ_2}	-	-	-	-	-	-	2.00	1.23	0.111
-2ℓ	3477.0			3475.1			3470.6		

[a]Description of dropout model parameter estimates: ρ_{γ_0} standard deviation of the random effect for the intercept and ρ_{γ_2} is the standard deviation of the random effect for the time

2007, pp. 386-387)

$$\mathrm{CD}_i = 2\left(\ell(\hat{\boldsymbol{\varphi}}) - \ell_{(-i)}(\hat{\boldsymbol{\varphi}})\right). \tag{6.7}$$

In this paper, we propose to study the influence of a potentially influential subject by measuring the discrepancy between the maximum likelihood ℓ^{nmar} and $\ell^{\mathrm{nmar}}_{(-i)}$ for the NMAR model, where $\hat{\ell}^{\mathrm{nmar}}_{(-i)}$ is the maximum likelihood estimate after the ith subject has been removed. We therefore propose to use the expression

$$\mathrm{CD}_i = 2\left(\hat{\ell}^{\mathrm{nmar}}(\hat{\boldsymbol{\varphi}}) - \hat{\ell}^{\mathrm{nmar}}_{(-i)}(\hat{\boldsymbol{\varphi}})\right) \tag{6.8}$$

as an indication of which subject is influential and is likely to derive the MAR missing data mechanism towards NMAR. The statistic (6.7) estimates the difference between the deviances for a model with all the subjects versus the same model with the ith subject removed, whereas the statistic (6.8) estimates the difference between the deviances for NMAR model with all the subjects versus the ith subject removed under the same NMAR model.

For the NMAR models (under Scenarios 1 and 2), we displayed only the index plots for the Scenario 1 in Figure 6.3, since the same patients are also influentials in the Scenario 2. The subjects 51, 129, 132, 136, and 137 have the largest CD_i for the NMAR models considered. This means that these subjects are potentially influential subjects for estimation of the parameters in the model. We observed that these subjects were randomized into the prednisolone arm. These subjects are

relatively influential for the following reasons. For example, the subject 51's large influence is caused by the very low (156) CD4 count level at baseline which suddenly increases to 187 (at 0.5 months) before dropout at 1 month visit. The subject 129's large influence is caused by the subject's very weak response (from 208 to 159 CD4 count level) to treatment. The subject dropped out after two weeks. The large influence of the subject 132 is caused by a very weak and constant CD4 count level (164) from baseline to 0.5 months. Also, the subjects 136 and 137's large influence is dues to a very weak and constant CD4 count level during their stay in trial. These subjects were removed in the subsequent analyses to study their effect on the parameter estimates and the model conclusions such as the dropout mechanism.

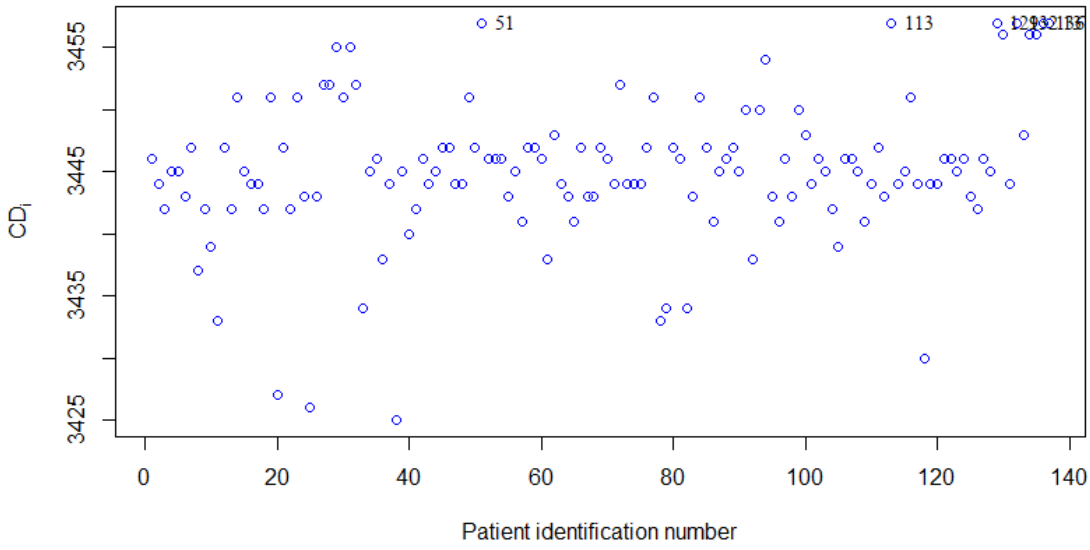

Figure 6.3: Index plots of CD_i for NMAR model with random intercept and slope shared between the measurement model and the dropout model.

Table 6.6 presents the results obtained after the potentially influential subjects (51, 129, 132, 136, and 137) have been removed from the analyses. The parameter estimates for treatment effect have increased relative to that in the Table 6.5. However, this increase is not statistically significant, as we saw before. Also, the parameter estimates do not change significantly under each missing data assumption. The results from the analyses considered in this paper suggest that the mechanism that generated the missing data in the IMPI trial is likely to be missing at random (MAR) mechanism. This conclusion is based on the following observations. For the different values of the sensitivity parameter $\gamma_k = k = [-10, 10]$, statistical inferences on parameter estimates (under NMAR, that is when $k \neq 0$) remain unchanged from that of the primary MAR analysis model (when $k = 0$). In addition, the log-likelihood value for the Scenario 1 (which assumes that the data are MAR) in Tables 6.5 and 6.4 is the highest suggesting that models that assume that the CD4 count values are MAR are more likely to provide best fits for the data. Also, statistical inferences on the parameter

Table 6.6: Parameter estimates for different missingness models fitted to the CD4 count data with subjects 51, 129, 132, 136, and 137 removed.

Mode	Scenario 1			Scenario 2			Scenario 3		
Parameter	Est.	s.e	p-value	Est.	s.e	p-value	Est.	s.e	p-value
Measurement model									
Time	0.41	0.12	< 0.001	0.49	0.14	< 0.001	0.47	0.15	< 0.002
Prednisolone	-0.12	0.83	0.888	-0.12	0.83	0.889	-0.174	0.83	0.835
Prednisolone × Time	-0.05	0.17	0.762	-0.03	0.17	0.883	-0.02	0.19	0.922
ART	2.66	0.57	< 0.001	2.61	0.57	< 0.001	2.60	0.57	< 0.001
Prednisolone × ART	-0.30	0.77	0.703	-0.29	0.77	0.706	-0.35	0.77	0.651
Age	3.00	1.49	0.046	3.01	1.49	0.045	-2.98	1.42	0.038
Dropout model									
γ_0	-	-	-	0.01	0.06	0.837	0.17	0.09	0.055
ρ_{γ_0}	-	-	-	1.57	1.07	0.144	0.95	1.34	0.481
γ_2	-	-	-	-	-	-	0.16	0.10	0.101
ρ_{γ_2}	-	-	-	-	-	-	2.30	1.30	0.079
-2ℓ		3328.0			3326			3312	

estimates under the Scenarios 2 and 3 (NMAR assumptions) are comparable to those under the Scenario 1 (MAR assumption). Since all the NMAR assumptions (that deviate from the MAR assumption) about the dropout process do not alter the statistical inferences and conclusion about the dropout mechanism, one can conclude the dropout mechanism is likely to be MAR.

6.5 Simulation study

In this section, we conduct simulation studies to evaluate the performance of the generalized shared-parameter model (GSPM) under different rates of missing data and mis-specification of the GSPM. We seek to investigate the ability of the GSPM to produce unbiased statistical inferences under various missing data rates. Also, we investigate how sensitive the GSPM is when mis-specified. This simulation is not intended to mimic any particular clinical setting. However, we used input parameters from the IMPI trial. The unstructured variance-correlation matrix was used in the simulation study. This variance-correlation matrix is common for both prednisolone and placebo groups. We generated a response at five time points.

The simulation study was carried out according to the linear mixed effect model defined by

$$\begin{aligned} \mathrm{Y}_{ij} &= \beta_0 + \beta_1 \times \mathrm{treatment}_i + \beta_2 \times \mathrm{time}_j \\ &\quad + \beta_3 \times \mathrm{treatment} \times \mathrm{time}_{ij} + \mathrm{b}_i + \epsilon_{ij}. \end{aligned}$$

In this simulations, initial values/parameters were obtained by fitting linear mixed effects model to CD4 count data from the IMPI trial. The initial values for $\beta_0, \beta_1, \beta_2$, and β_3 are 13, 0.75, 0.11, -0.19, 0.20 respectively. The initial value for standard deviation σ of the random effect b_i is 4.57 and that of the ϵ_i is 3.07. In order to generate the data sets, we assumed that (1) the measurement at the first time point ($j = 0$) from the original data set is completely observed, (2) the data are MAR or NMAR mechanism, (3) the missing pattern is monotone, and (4) there are different dropout rates. The probability of a missing value of Y_{ij} is $\Pr(R_{ij} = 1) = \Phi(\vartheta_0 + \vartheta_1 Y_{i,j-1})$ for MAR mechanism and $\Pr(R_{ij} = 1) = \Phi(\vartheta_0 + \vartheta_1 Y_{i,j-1} + \vartheta_2 Y_{ij})$ for not missing at random. We chose ϑ_0 and ϑ_1 so that the amount of missing data in the datasets are approximately 5%, 20%, and 50% for MAR mechanism and ϑ_0, ϑ_1, and ϑ_0 and ϑ_2 so that the missing data in the dataset are approximately 5%, 20%, and 50% for NMAR mechanism.

We considered two steps, called M-step and D-step, for generating the simulated data sets. The longitudinal measurements are generated under the M-step and the MAR and NMAR mechanisms are simulated under the D-step.

M-step: In the IMPI trial, prednisolone effect was not significant, so in this simulation study, we assume that the prednisolone effect is significant. We generated five-repeated measurements for each patient by a random number from a multivariate normal distribution. These processes were repeated 1000 times for 200 patients. These patients were randomly assigned to two treatment (active and placebo) groups in a ratio 1:1 (Iddrisu and Gumedze, 2019*a*).

D-step: In this step, we generated missing data under the MAR and NMAR mechanisms. These missing data were generated using a logistic regression model describing the probability of dropout. For the MAR mechanism, we generated the missing data by dropping observations according to a logistic regression model relating the probability of dropout at particular time point with changes from baseline to previous time point. For the NMAR mechanism, missing data were simulated by dropping observations according to a logistic regression model describing the probability of dropout at particular time point with changes from baseline to previous and current time points (Iddrisu and Gumedze, 2019*a*). We chose specific values for the logistic regression model in order to produce the desire dropout rates in a given missing data mechanism. Under each of the missing data mechanisms, we generated overall dropout rates at 5%, 30, and 50%.

The sub-models M_1-M_9 were then implemented to the imputed data sets (under the three missingness rates 5%, 30%, and 50%) using the SAS procedure NLMIXED (discussed in Section 6.3). We evaluated the performance of these sub-models against the usual multiple imputation (MI) approach for handling missing data. The MI is known to provide valid statistical inferences when missing values are missing at random (MAR) (Rubin, 1996, 1976; Ayele et al., 2014; Carpenter et al., 2013; Molenberghs et al., 2008).

The simulation results showed that the estimates for treatment effect is significant. These estimates further suggest that there is a significant reduction in the response for patients in the treatment arm, relative to patients in the placebo. Our focus is to evaluate the performance of the sub-models M_1-M_9 against the MI approach by computing the bias, mean square error (MSE), and coverage probability (CP) for the treatment effect. We found that the sub-models, M_1-M_9, produced comparable estimates for the model performance indicators (bias, MSE, and CP). Hence we presented only the results of the sub-model M_6 versus the MI and MGSPM (6.9) approaches in Table 6.7.

The results showed that the parameter estimates are unbiased with relatively lower MSE and the coverage probabilities are approximately 95% for the MAR mechanism. As expected, the MI and MGSPM approaches produced biased estimates of treatment effect under the NMAR mechanism. This is because the MI approach assumes that the data are missing at random whereas the MGSPM is a mis-specified form of the GSPM. The simulations results showed the GSPM is robust to the amount of missing data in a study since the M_6 produces unbiased estimates with relatively very low RMSE values (even when the missing data rate is high as 50%) and coverage probabilities are approximately 95%. These results suggest that GSPM is sensitive to mis-specification, especially when data are assumed to be NMAR mechanism

Table 6.7: Simulation results under MAR and NMAR mechanisms by missingness rate

missingness rate		MAR			NMAR		
5%	Analysis	Bias	RMSE	CP	Bias	RMSE	CP
	MI	0.135	0.018	95.20%	2.871	1.342	99.40%
	M_6	0.149	0.022	95.10%	0.134	0.018	95.30%
	MGSPM	0.292	0.092	96.60%	4.123	3.831	99.60%
30%							
	MI	0.262	0.069	95.10%	3.121	2.813	99.00%
	M_6	0.283	0.080	95.30%	0.178	0.078	95.50%
	MGSPM	0.356	0.123	97.40%	5.213	3.876	98.90%
50%							
	MI	0.273	0.075	95.40	6.821	4.345	99.90
	M_6	0.292	0.085	95.60%	0.123	0.0643	95.70%
	MGSPM	0.532	0.245	98.50%	7.452	4.352	99.20%

As part of the simulation studies, we further assessed how sensitivity of the GSPM approach is when it is mis-specified. For instance, we assessed sensitivity of a mis-specified form of a GSPM (MGSPM) as

$$\Pr\left(\mathbf{Y}_i, \mathbf{R}_i \mid \mathbf{g}_i, \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(\mathbf{Y}_i^o\right) \Pr\left(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o, \mathbf{k}_i\right) \Pr\left(\mathbf{R}_i \mid \mathbf{k}_i\right). \tag{6.9}$$

The MAR counterpart defined as

$$\Pr\left(\mathbf{Y}_i, \mathbf{R}_i \mid \boldsymbol{\theta}, \boldsymbol{\psi}\right) = \Pr\left(\mathbf{Y}_i^o\right) \Pr\left(\mathbf{Y}_i^m \mid \mathbf{Y}_i^o\right) \Pr\left(\mathbf{R}_i\right)$$

This model is a mis-specified form of the GSPM (MGSPM) because it is assumed that the the random effects within the observed sequence are assumed independent which is no longer a GSPM or a joint model for non-ignorable missing data. It follows that

$$\mathrm{E}[\mathbf{Y}^o \mid \mathbf{b}_i] = \beta_0 + k_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i$$

$$\mathrm{E}[\mathbf{Y}^m \mid \mathbf{Y}^o, \mathbf{b}_i] = \beta_0 + k_i + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j + \beta_4 a_{ij} + \beta_5 a_{ij} T_i + \beta_6 c_i.$$

The corresponding model for the probability dropout is specified as

$$\begin{aligned}\operatorname{logit}\left[\Pr\left(R_{ij} = 1 \mid R_{i,j-1} = 0, \mathbf{b}_i, T_i, t_j, \gamma\right)\right] = \gamma_0 + \gamma_k k_i + \gamma_1 T_i + \gamma_2 t_j + \gamma_3 T_i t_j \\ + \gamma_4 a_{ij} + \gamma_5 a_{ij} T_i + \gamma_6 c_i\end{aligned}$$

The parameter estimates from MGSPM are presented in Table 6.8. Compared with the results in the Table 6.4, it can be observed that the parameter estimates under the Table 6.8 and the other sub-models (M_1-M_9), are different in terms of sign and magnitude. In addition, the MGSPM produces the highest estimate for the residual variance relative the sub-models.

6.6 Discussion and conclusion

When data are missing, any method of analysis should take into account the missing data process in order to draw valid statistical inferences. The missing data process are classified into MCAR, MAR, and MNAR mechanisms, and modeling frameworks SeM, PMM, and GSPM, can be used to handle the missing data mechanisms. One should pay attention when fitting an MNAR model, regardless of framework, because they rest on strong, unverifiable assumptions (Molenberghs et al., 2008). Therefore, a sensitivity analysis should be part of the data analysis, especially when dealing with missing data. Various authors proposed sensitivity analysis methods for the SeM framework and PMM. A sensitivity analysis for the GSPM framework has recently been considered Creemers et al. (2010). In this paper, we applied this sensitivity analysis tool for CD4 count data from the IMPI trial. This is done by letting the values for unidentifiable sensitivity parameters take all values

Table 6.8: CD4 count data: parameter estimates, standard errors of the MGSPM fits cont

Effects	Par	MGSPM
Mean structure		
Measurement		
prednisolone	β_1	0.19 (0.619)
Month	β_2	0.31 (0.171)
prednisolone × Month	β_3	0.14 (0.235)
ART	β_4	1.96 (0.716)
Prednisolone × ART	β_5	-0.02 (0.985)
Age	β_6	-3.29 (0.842)
Dropout		
Prednisolone	γ_1	1.06 (1.683)
Month	γ_2	1.82 (0.468)
Prednisolone × Time	γ_3	-0.37 (0.401)
ART	γ_4	-5.56 (1.837)
Prednisolone × ART	γ_5	1.77 (2.012)
Age	γ_6	-0.04 (2.203)
Variance-covariance structure		
Measurement		
Resid.var.	σ^2	27.51 (1.602)
Rand.int.var.	ρ_g^2	
	ρ_h^2	
	ρ_k^2	28.30 (13.473)
	ρ_j^2	
Dropout		
Scale factor	γ_g	
	γ_k	Fixed: $\gamma_k = 1$
	γ_j	
Rand.int.var.	$\gamma_g^2\rho_g^2$	
	$\gamma_j^2\rho_j^2$	

over a pre-specified grid. Based on empirical Bayes estimates, the missing values in the original data set are imputed a number of times. Using multiple imputation tools, these data sets are analyzed and inferences obtained. We also proposed to study the impact of potentially influential subjects, on parameter estimates and dropout mechanism, using global influence approach for the

GSPM. If results are very sensitive to the value of the sensitivity parameter, inferences should not be considered trustworthy.

The results showed that there is no significant prednisolone, prednisolone × time, prednisolone × ART effects and the effects of time, ART, and age are significant. The empirical Bayesian ideas are used in this paper. However, sensitivity analysis based on a fully Bayesian ideas has not been presented here. Although such approach would be possible but falls outside the scope of this manuscript.

In this paper, the sensitivity analyses approaches based on the GSPM Creemers et al. (2011) were applied to monotone data. However, these sensitivity analyses approaches can be applied to non-monotone data Gao (2004); Tsonaka et al. (2009). Also, the sensitivity analyses approaches in this paper were applied to continuous longitudinal measurements. However, these approaches can also be applied in incomplete longitudinal binary data settings (Gao, 2004). In this study we assessed the importance of conducting some forms of sensitivity analysis and applied principles in the IMPI trial (Mayosi et al., 2014) setting. The IMPI trial was a cardiology trial in which HIV data were relevant and collected. However, the HIV data were not collected as they would have been in a HIV-related clinical trial, and hence there are missing CD4 counts. Although the IMPI trial was a cardiology trial, our analyses of the HIV data provide reasonable information regarding the degree of influence of ART and prednisolone treatments on CD4 changes over time.

The missing values of the CD4 count might be informative, and hence later values of CD4 count are missing because 246 (18%) of subjects have died. This would require joint modeling on the CD4 count and time to death. Various authors (Baghfalaki et al., 2017; De Gruttola and Tu, 1994; Rizopoulos et al., 2008; Little, 1995) have proposed such joint models (JM) within the shared-parameter model framework. Viviani et al. (2014) considered sensitivity analysis to the non-ignorability of the dropout process in the JM. Their sensitivity analysis approach was based on the Index of Local Sensitivity to Non-Ignorability (ISNI) proposed by Troxel et al. (2004). The ISNI evaluates the rate of change of parameter estimates to the assumed degree of non-ignorability in the neighborhood of an ignorable model (Xie et al., 2018). Viviani et al. (2014) extended the ISNI proposed by Troxel et al. (2004) to joint models. Xie et al. (2018) have recently developed a new **R** package, known as **isni**, to conduct such sensitivity analysis. A discussion on limitations and recommendations of the **isni** package can be found in Xie et al. (2018). One of the limitations is that the **isni** package cannot accommodate missing covariates and hence uses *ad hoc* means to process the data for analysis when covariates have missing values (Xie et al., 2018).

Chapter 7

Discussion and conclusions

7.1 Summary and sensitivity analysis methods

The effect of medical treatment can be appropriately evaluated via randomized longitudinal studies designs. Longitudinal studies are designed to repeatedly record measurements on a response variable of interest for each patient at some selected scheduled visit. However, such designs are often plagued by missing data. This is because clinicians or researchers are often not able take measurements at all the scheduled visits and the occurrence of missing data is often beyond the control of the researcher. Missing data may be caused by early withdrawals or noncompliance due to the adverse effect of trial's interventions or protocol deviations. Depending on the study protocol, such deviations may be defined as poor compliance with the intervention or withdrawal from the intervention, unblinding either of the intervention or evaluation and moving to partial compliance with treatment or dropout such that no further information is recorded on the patient (Carpenter et al., 2013).

Missing data in clinical trials can undermine the benefits provided by randomization into placebo-treated and active-treated groups (NRC, 2010; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013). Incomplete longitudinal data leads to loss of information and a reduction in the precision with which longitudinal change can be estimated (Molenberghs and Kenward, 2007; Fitzmaurice et al., 2008). The presence of missing data may complicate analysis of the incomplete longitudinal data and interpretation of the results. Two approaches to the problem of missing data are to reduce/minimize the frequency of missing data in the first place and to use appropriate statistical techniques that account for the missing data. The former approach is preferred, since the choice of statistical method requires unverifiable assumptions concerning the mechanism that generates the missing data, and so always involves some degree of subjectivity (NRC, 2010). This means that valid statistical inferences can be obtained under such analysis by incorporating the reasons for the missing data in the statistical analysis (Rubin, 1976; Little, 1995). These reasons are often referred to as missing data mechanisms (Little, 1995; Rubin, 1976).

Analyses, such as complete-case analyses, that do not account for the missing data or assume that the data are missing at random (MAR) are likely to produce biased statistical inferences, especially when the data are not missing at random (NMAR) (Molenberghs and Kenward, 2007). This means that to draw valid statistical inferences, methods of analysis should be based on plausible scientific

assumptions about the missing data. This requires sensitivity analyses to investigate the robustness of statistical inferences to plausible alternative assumptions about the missing data. Specifically, there is the need to perform sensitivity analysis to assess the sensitivity of statistical inferences under the MAR primary analyses to alternative plausible assumptions under the NMAR analyses.

In this thesis, the aim is to conduct sensitivity analysis to investigate the robustness of statistical inferences under MAR assumption to alternative assumptions under NMAR. This thesis also investigated the effects of prednisolone and prednisolone × ART interaction on patients' CD4 counts evolution. We decided on 1a suitable primary analysis method for the incomplete CD4 count data from the IMPI trial and then assessed the robustness of statistical inferences, under the primary analysis, to alternative assumptions about the missing data through sensitivity analysis. In this thesis, our primary analysis methods assumes that the data are MAR. The NMAR sensitivity analyses methods considered are based on the selection (Diggle and Kenward, 1994; Verbeke et al., 2001; Kenward, 1998), pattern-mixture (Carpenter et al., 2013; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013), and shared-parameter (Rubin, 1976; Wu and Bailey, 1988; Wu and Carroll, 1988; Gao, 2004; Tsonaka et al., 2009; Creemers et al., 2010, 2011) and model frameworks.

In Chapter 4, our sensitivity analysis to nonrandom dropout is based on the Diggle and Kenward (1994) selection modeling framework. The DK selection model factorizes the joint distribution of the measurement and the missingness mechanisms into marginal distribution of the measurement process and the conditional distribution of the missingness process, given the response. We assumed the linear mixed effect model for the measurement process. The probability of dropout at a given scheduled visit time j was modeled using a logistic regression model. The DK model is specified in such a way that the response variable in the measurement model is a covariate in the dropout model. The measurement and the dropout models are linked in the dropout model by the parameter estimate ψ_2. This feature of the selection model makes it more convenient for some form of sensitivity analyses (Diggle and Kenward, 1994).

For instance, whenever $\psi_2 = 0$, the missing data mechanism is assumed to be missing at random (MAR). This implies that dropout no longer depends on the current measurement Y_{ij} but on previous observation $Y_{i,j-1}$. On the other hand, whenever $\psi_2 \neq 0$, the missing data mechanism is NMAR. This assumption has implication for sensitivity analysis to nonrandom dropout (NMAR). For instance, under such an assumption, one is able to investigate how sensitive the statistical inferences under MAR mechanism are to various alternative assumptions under a particular NMAR model considered.

However, the DK selection model (Diggle and Kenward, 1994) received several criticisms ranging from computational instability (Shen et al., 2006) and the use of such models (Thijs et al., 2000).

In particular, conclusions based on such models have been questioned in terms of their reliability because of their strong but untestable assumptions about the NMAR mechanism (Molenberghs and Kenward, 2007). The nature of the incompleteness was questioned since it was due to the manner in which the study was designed (Verbyla and Cullis, 1990) but not a genuine dropout (Kenward, 1998; Thijs et al., 2000).

Diggle and Kenward (1994) applied their selection model to the milk protein data (Verbyla and Cullis, 1990) and concluded that the dropout mechanism was nonrandom. However, various authors have argued that such models should not be used to conclusively determine whether or not the dropout mechanism is nonrandom. For instance; (1) Little (1995) argued that parameter estimates depend on normal assumptions and correct specification of the model, (2) Laird's discussion to Diggle and Kenward (1994) noted that estimating the untestable assumptions can only be achieved by making modeling assumptions about the dropout mechanism. Thus, the consequences of model misspecification will probably be far more severe when the dropout mechanism is non-ignorable (when the missing data cannot be ignored), (3) Rubin, in his discussion to Diggle and Kenward (1994) selection model, argued that inferences for the data parameters generally depend on the assumed missingness mechanism; which implies greater sensitivity of inference to a reasonable model specification, and (4) Molenberghs et al. (1997) argued that conclusions are conditional on the appropriateness of the assumed model; which is untestable. It was finally agreed that such analyses should be placed within sensitivity analysis framework.

Consequently, Verbeke and Molenberghs (2009) conducted sensitivity analysis based on the DK selection modeling framework. They showed that excluding a small number of measurement errors considerably changes the likelihood ratio test statistics for the MAR null hypothesis. Kenward (1998) re-analyzed the milk protein data dataset using the DK selection model. He demonstrated that removing two unusual observations from the analyses changed the nonrandom dropout mechanism towards random dropout mechanism. This implies that the presence of these two subjects in Diggle and Kenward (1994) analyses lead to the conclusion that the dropout mechanism was nonrandom. Molenberghs, Verbeke, Thijs, Lesaffre and Kenward (2001) confirmed Kenward (1998) results using influence diagnostics. These considerations motivated the use of the influence diagnostics (Cook, 1986; Verbeke et al., 2001; Shen et al., 2006; Thijs et al., 2000) to conduct sensitivity analysis to the incomplete CD4 counts measurements from the IMPI trial (Mayosi et al., 2014). This is because the unobserved CD4 counts measurements in the IMPI clinical trial were due to administrative reasons but not genuine dropout as in the case of the of milk protein clinical trial (Verbyla and Cullis, 1990). Hence, it is unlikely that the probability of dropout will depend on the unobserved CD4 counts measurements.

We also considered the local influence approach for sensitivity analysis under the DK selection model. The local influence approach (Verbeke et al., 2001) allows one to assess the influence of

potentially influential subjects on parameter estimates and drop out mechanism by perturbing the MAR model in the direction of NMAR model. This would be achieved by perturbing the NMAR part ψ_2 of the DK selection model. This approach allows one patient to dropout randomly while the remaining patients dropout non-randomly (Fitzmaurice et al., 2008; Molenberghs et al., 2014). Thereafter, potentially influential patients identified were removed in the subsequent analyses to assess their impact on the estimation of the parameter estimates in the models.

One important property of the local influence approach is that one is able to attribute the influence of potentially influential patients to their specific characteristics by inspecting the index plots components of the model (Lesaffre and Verbeke, 1998; Verbeke et al., 2001; Shen et al., 2006). That is, by removing patients and comparing results before and after, you may see changes in the missing-data mechanism model. This may then explain that there is influence by the potentially influential patients on the missing data mechanism. In this study, the components inspected include (1) the fixed effects, (2) variance of the random effects, (3) parameters describing the dropout mechanism (4) the maximum influence measure and (5) total influence. The total influence is the overall influence from components 1, 2 and 3.

We also considered a stress-testing methodology, under the DK selection model, to assess the impact of lowering or increasing dropout probability on treatment effect and key conclusions. We assessed the impact of dropout at the respective treatment arms, on the treatment effect, using separate dropout models for each treatment arm. Fitting separate dropout models for each arm allows for different departures from MAR for prednisolone and placebo arms. This also allowed us to determine the retention rate in each arm, which is likely to overturn treatment effect and conclusions and may have clinical interpretation or meaningless.

In Chapter 5, we considered an alternative sensitivity analysis approach known as the pattern-mixture with multiple imputation approach (PM-MI) (Carpenter et al., 2013). This approach is based on the conventional pattern-mixture model (PMM). The PMM is a reverse factorization of the selection model (Little, 1995; Rubin, 1976) considered in the Chapter 4. The key idea in the PM-MI approach is a "reference based" imputation. Instead of specifying parameter values describing how post-deviation behavior evolves, the PM-MI approach identifies groups of patients within the trial with this kind of behavior, and then uses this to construct the imputation distribution (Carpenter et al., 2007, 2013; Mallinckrodt, Lin and Molenberghs, 2013; Ayele et al., 2014). This means that for each patient who deviates from the study protocol or drops out, the joint distribution of their pare-deviation and post-deviation behavior is built. The joint distribution uses information from other groups of patients in the trial (the reference group) and then use this information to construct the deviated patients' conditional distribution of post-deviation given pre-deviation data. This conditional distribution of post-deviation data given the pre-deviation data is in turn used to impute deviated patients' missing post-deviation data (Carpenter et al., 2013; Ratitch et al., 2013).

Under the PM-MI approach, two estimands of interest are the *de jure* and the *de facto* estimands (Carpenter et al., 2013). The *de jure* estimand estimates the effect of the treatment among patients who essentially adhered to the study protocol whereas the de facto estimand estimates the effect of treatment, regardless of whether patients adhered to the study protocol. Primary analysis was based on the de jure estimand and sensitivity analysis was based on the de facto estimand hypotheses about the missing data. The de jure estimand was based on the multiple imputation (MI) under MAR mechanism whereas the sensitivity analysis was based on the four de facto estimand hypotheses about NMAR mechanism (Carpenter et al., 2013).

Whereas these assumptions about the missing post-deviation outcome data have been proven to be practical and permits relevant, accessible assumptions for framing primary and sensitivity analyses, the approach depends on the relevance of the assumptions about missing post-deviation data in relation to the context of the trial at hand. For instance, in the IMPI trial, dropouts from one arm can assume response distribution of patients in the other arm. This is because, patients in treatments are expected to have comparable treatment effect or the same statistical behavior (Carpenter et al., 2013; Shen et al., 2006; Iddrisu and Gumedze, 2018, 2019*a*,*b*). However, in trial settings where one arm is expected to have low treatment effect compared with the other arm, then one will have to adopt methods that will reduce treatment effect in such arm by a certain amount (Ratitch et al., 2013).

The PM-MI was implemented to the incomplete CD4 count measurements from the IMPI trial to obtained K complete data sets. The linear mixed effect model was fitted to each of the completed data sets. Parameter estimates were combined using the Rubin's rule. The combined parameter estimates with their associated standard errors were then used for the final statistical inferences.

We evaluated the performance of the PM-MI approach (LMCF, J2R, CR, and CDR) under different rates of missing data and investigate whether these methods will be able to produce unbiased statistical inferences under different missingness rates. The performance of the PM-MI approach was evaluated using the bias, mean square error (MSE), and the coverage probability of the parameter estimates of interest. We found that the J2R and CR methods are robust to higher rate of missing values and are highly recommended when dealing with data with substantial amount of missing data. However, the CDR and LMCF methods can also be used when the amount of missing data is less than 50%.

In Chapter 6, we implemented an alternative sensitivity analysis approaches based on the shared-parameter model (SPM) framework (Rubin, 1976; Wu and Bailey, 1988; Wu and Carroll, 1988; Gao, 2004; Tsonaka et al., 2009; Creemers et al., 2010, 2011). The SPM, by specification, is a joint model for non-ignorable missing data (Tsonaka et al., 2009; Gao, 2004). One important feature of the approach is that it assumes that the measurement and the dropout models shared

a set of random effects (Roy, 2003; Mallinckrodt, Lin and Molenberghs, 2013; Albert et al., 2002), where the random effects account for the association between the measurement and the missingness models. These random effects also account for the correlation between the repeated measurements. The random effects are assumed to be normally distributed with zero mean and a given standard deviation $\boldsymbol{\sigma}$. A logistic regression model is assumed for the probability of dropout. The $\boldsymbol{\sigma}$ is the association parameter between the dropout model and the measurement model. The dependence of the measurement model and the dropout model is determined using the standard deviations $\boldsymbol{\sigma}$ of the random effects. The magnitude of this standard deviation determines the extent of association between the measurement and the dropout models as well as the level to which inferences are likely to be affected by such dependence.

In the context of the IMPI trial, our initial analyses of the incomplete longitudinal CD4 count data revealed that there is variability associated with the random intercept. This means that each patient draws a random slope from a normal distribution with mean zero and standard deviation. Since this random slope governs the patient's expected rate of decline in the outcome and the probability of dropping out (see Molenberghs and Fitzmaurice in Fitzmaurice et al., 2008, Ch. 19), it is important consider a model for the dropout probability such that the probability of dropout depends on the random effects rather than the response (as seen in DK selection model in Chapter 4, Section 4.2.2). The SPM framework provides a tool for joint modeling of the measurement and the dropout processes, where the probability of dropout depends on the random effects. So here it is assumed that that the measurement process model (mixed effects model) and the dropout process model share a set of random effects. In this way, we will be able to make different assumptions regarding the association between the measurement model and the dropout model (via the random effects), and then compare the results to assess the robustness of the statistical inferences to these assumptions.

It follows that whenever $\boldsymbol{\sigma} = \mathbf{0}$ (assuming MAR mechanism), then the measurement and the dropout models no longer share common random effects. Thus, valid inferences can be obtained by fitting the two sub-models separately using standard methods of analysis. On the other hand, whenever $\boldsymbol{\sigma} \neq \mathbf{0}$, the missingness mechanism is described as non-ignorable, where one cannot no longer ignores the missing data mechanism and fit the models separately to obtain valid inferences (NMAR). This feature of the shared-parameter model has implication for assessing sensitivity of results under MAR to alternative plausible assumptions under NMAR models. For instance, one can conduct analysis under the assumption that $\boldsymbol{\sigma} = \mathbf{0}$ and under the assumptions that $\boldsymbol{\sigma} \neq \mathbf{0}$ and then compare the results.

In this chapter we focused on the general SPM (GSPM) (Creemers et al., 2010), which is the most general form of the SPM. The literature has so far focused on selection model (Diggle and Kenward, 1994), pattern-mixture model (Little, 1993), and shared-parameter model (Follmann and

Wu, 1995*a*) frameworks for the analysis of a longitudinal measurements with nonrandom dropout process. However, not so much has been done for the shared-parameter model framework (Viviani et al., 2014) except the paper by Creemers et al. (2011).

We have implemented the models under the GSPM and conducted sensitivity analysis to the incomplete longitudinal CD4 count data. Here, we implemented the sensitivity analysis approach considered in Creemers et al. (2011) paper and then proposed to assess the effect of potentially influential influential subjects, on parameter estimates and dropout mechanism, using the global influence approach. We further conducted simulation studies to evaluate the performance of the GSPM under different rates of missing data and mis-specification of the GSPM. Specifically, we investigated the ability of the GSPM to produce unbiased statistical inferences under various missing data rates and also, how sensitive the GSPM is when mis-specified. The simulation results revealed that the GSPM is robust to different rates of missing data and sensitive to mis-specification (see Iddrisu and Gumedze (2019*b*) and Chapter 6 of this thesis).

The three model frameworks summarized in this section are similar in the following ways: Each of these frameworks represents a joint model for the measurement and the dropout mechanism. Also, each framework can be characterized according to the three missing data mechanisms (MCAR, MAR, NMAR) (Molenberghs et al., 2008; Creemers et al., 2010). More importantly, the specification of each framework allows for conducting some form of sensitivity analysis to missing data assumptions. In addition, multiple imputation can be considered under each of these model frameworks. For example, see Smuk (2015); Rezvan et al. (2015) for the selection model framework, Carpenter et al. (2013); Ayele et al. (2014); Mallinckrodt, Lin and Molenberghs (2013); Ratitch et al. (2013) for the pattern-mixture model framework, and Creemers et al. (2011) for the shared-parameter model framework.

However, these model frameworks differed in the manner in which the joint model is specified. For instance, in the selection model (Diggle and Kenward, 1994), the measurement model and the dropout model are link in the dropout model (by the ψ_2) in such a manner that the response variable $\mathbf{Y}_i$ in the measurement model is covariate in the dropout model. The pattern-mixture model (Little, 1993) is a reverse factorization of the selection model and is specified in such a way that the observed data is a mixture of dropout patterns weighted by their respective probability of missingness. Unlike the selection model framework, which assumes that the dropout probability depends on the response variable, the shared-parameter model framework (Albert and Follmann, 2000; Follmann and Wu, 1995*a*; Creemers et al., 2010) assumes that dropout probability depends on a set of common random effects.

7.2 Results

Analysis of the CD4 count data from the IMPI trial (Mayosi et al., 2012, 2014) revealed that the time, age, and ART effects are statistically significant (see Tables 4.1, 4.2, 5.1, 5.2, and 6.4). There is an observed increase in the level of patients CD4 count over time and subjects received ART at each schedule visit have higher CD4 count levels relative to those who are not on ART treatment at each schedule visit. The results also showed that older patients are more likely to have lower CD4 count than younger patients. The main treatment (prednisolone), and the prednisolone $\times$ ART and prednisolone $\times$ time interactions effects are not significant. The implication is that prednisolone treatment does not influence the ART treatment. We also found a reduction of CD4 count level in the prednisolone arm. This finding agrees with the Figures 2.2 and 2.4, where this reduction is observed after the 3 months visit. However, this reduction is not significant.

The near zero estimates of the prednisolone $\times$ time interaction effect (see Tables 4.1, 4.2, 5.1, 5.2, and 6.4) suggest that there is no difference in prednisolone effect in both arms over time. This means that the effect of treatments in both arms does not differ significantly over time. The local influence results showed that removal of patients identified as potentially influential overturned the effects of the parameter estimates and conclusion about the dropout mechanism. However, the removal of these influential patients from the analyses leads to bias prednisolone effect. This gives an indication that valid analysis must include all the potentially influential patients. The stress-testing (see Figure 4.3) results revealed significant prednisolone effect for some combinations of dropout rates, which may have clinically interpretation or meaningless. These significant prednisolone effects give an indication that the prednisolone effect in the IMPI trial would have been significant if we could ensure that the dropout probability caused by a positive change on the current outcome values decreases in the prednisolone arm and increases in the placebo arm. The global influence (see Table 6.6) results agree with the results under the local influence in a sense that the effect of the removal of the influential patients produced biased prednisolone effect, suggesting that all patients should be considered in the analyses in order to produce unbiased statistical inferences.

Our simulation studies (see Section 5.6 and results in Appendix B.2) to evaluate the performance of the LMCF, J2R, CDR and CR methods, under different rates of missing data, revealed that these method produced unbiased estimates for the parameter of interest. The simulation results also revealed that, even when the missingness rate was substantial as 50%, the methods still produced unbiased estimates except LMCF and CDR methods (which produced less unbiased estimates for time effect). The LMCF method produces biased estimate for time effect because it does not account for variability due to the imputation of missing data and hence produces large number of identical imputed data as the amount of missing data increase. Also, because the CDR method assumes that after a patient deviates that patient's post-deviation mean increments copy those from the

reference arm, this increments are likely to be identical in the presence of high amount of missing data leading to biased effect of time.

The J2R and CR methods appeared to be robust to higher rate of missing values. The J2R method robustness can be attributed to its assumption that after a patient stops taking treatment from the randomized arm, such patient's mean response distribution is now considered to be the same as that of the "reference" group of patients since this assumption accounts for variability due imputation missing data. This explained why the CR method also appears to be robust to substantial amount of missing data since a patient's whole distribution, both pre-deviation and post-deviation data, is the same as reference arm. Our simulation results therefore suggest that the J2R and CR should be used when dealing with data with substantial amount of missing data. However, the CDR can also be used since a study with amount of missing value above 50% would be considered, in most studies, as extremely large. Thus, these methods (proposed by Carpenter et al. (2013)) can be used for handling the missing data in the IMPI clinical trial and other trials with similar settings. By definitions, the CR and J2R methods are recommended for used in the IMPI trial setting and other trials with similar settings because patients in either arm are expected to obtain equal benefit from ART.

The parameter estimates produced under the generalized shared-parameter model (GSPM) approach are always almost the same for the different values of the sensitivity parameter γ_k. This gives an indication about the level of robustness of statistical inferences under MAR analysis to alternative assumptions under the NMAR analyses. Also, we conducted simulation studies to evaluate the performance of the generalized shared-parameter model (GSPM) framework under different missing data rates and when the GSPM is mis-specified (MGSPM) (see Section 6.5 for studies design and results in Table 6.7). We found that the parameter estimates are unbiased with relatively lower MSE and the coverage probabilities are approximately 95% for the MAR mechanism. However, the MI and MGSPM approaches produced biased estimates of treatment effect under the NMAR mechanism. This is because the MI approach assumes that the data are missing at random whereas the MGSPM is a mis-specified form of the GSMP. The simulations results showed the GSPM is robust to the amount of missing data in a study since the M_6 produces unbiased estimates with relatively very low RMSE values (even when the missing data rate is high as 50%) and coverage probabilities are approximately 95%.

7.3 Remarks on the sensitivity analysis

We note that sensitivity analysis does not imply a way of refuting conclusions from a given primary analysis method. Rather, it is a way by which more confidence can be placed on the primary analysis

conclusions. When analysis with and without influential subjects produce the same conclusion, for instance, the trial's treatment effect, one will place more confidence in the conclusions than when no sensitivity analysis had been conducted (Jansen et al., 2006).

One major advantage of the selection models is that they directly modeled the marginal distribution of the complete data, which is often the usual focus of statistical inferences in longitudinal study. Statisticians prefer selection models since they offer straight forward formulation of the hypothesis about the non-response process (Diggle and Kenward, 1994; Diggle et al., 2002).

Although assumptions about the non-responses are clear in selection models, what is less clear is how these translate into assumptions about the distributions of the unobserved outcomes (Diggle and Kenward, 1994). That is, selections models are often under-identified and identification is achieved by assuming models for the dependence of missingness process on the unobserved outcomes. Unless in a very simple case, it can be difficult to determine a specific identifying restriction that must be placed on such models (Cole et al., 2005; Glonek, 1999). The selection models can also be computationally demanding.

The pattern-mixture models are straightforward to fit as standard models that assume that non-response is ignorable. The pattern-mixture models are often over-specified or under-identified such that parameters describing the incomplete patterns cannot be estimated (Carpenter et al., 2007; Carpenter and Kenward, 2012; Mallinckrodt, Lin and Molenberghs, 2013). Identification is achieved by assuming unverifiable links among the distribution of the outcomes conditional on the missingness patterns. It is straightforward to determine and identify restriction in selection models than in pattern-mixture models as the number of patterns increased (Little, 1993).

It can be observed that both the selection and the pattern-mixture models specifications contain parameters that cannot be estimated without imposing arbitrarily modeling assumptions about the distribution of such parameters or distributions describing the incomplete patterns or data (Carpenter et al., 2007, 2013).

The pattern-mixture approach with multiple imputation (PM-MI) is easily understood by non-statisticians since the distribution of the estimates from both the complete and imputed records can be displayed graphically as a starting point for discussion about NMAR mechanism. The PM-MI approach allows great flexibility in the specification of post-deviation distributions, with the same substantive model fitted to the imputed data. Another advantage of approach is that it is well suited to framing relevant, accessible assumptions for both primary and sensitivity analyses for clinical trials with longitudinal follow-up. Its focus on different response patterns for post-deviation data makes it accessible. That is, information on the model for post-deviation data can come from other covariates (including the stated reason for deviation), as well as external information. With the PM-MI approach, different groups of patients, and even if desired different patients, can

be modelled differently. In addition, the way this information will be used can be pre-specified, naturally presented graphically, and agreed upon with interested parties before the treatment code is broken.

One attraction of the PM-MI approach is that it does not involve stakeholders in specifying distributions of unknown parameters describing post-deviation profiles, such as the difference between mean responses post-deviation, or the difference in the mean slope of response profiles post deviation.

Whenever the interest to built a model for longitudinal measurements and time to first occurrence of an event (survival analysis) such as death or disease elimination, the shared-parameter model framework provides a natural framework for implementation such models. The models build for this type of analysis are known as joint models (JM). Information on JM can be found in (De Gruttola and Tu, 1994; Little, 1995; Rizopoulos et al., 2008; Ali et al., 2011; Viviani et al., 2014; Baghfalaki et al., 2017) and sensitivity analysis for the JM can be found in (Viviani et al., 2014; Xie et al., 2018).

7.4 Thesis contribution

Firstly, the objective of every trial varies depending on the estimand of interest (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; NRC, 2010). This means that the validity of any method of analysis will depend on the trial setting (estimand of interest). In this thesis, we have considered sensitivity analyses that can be applied to the IMPI trial and other trials with similar settings.

Secondly, we have conducted sensitivity analyses to investigate the effects of various degree of dropout probabilities on treatment effect. These analyses provide a tool for medical and clinical researchers to stress-test missing data assumptions in order to assess the impact of such assumptions on the statistical inferences. The work on these sensitivity analyses and a SAS code have been published (Iddrisu and Gumedze, 2018).

Thirdly, Carpenter et al. (2013) proposed and applied the PM-MI approach to measurements at the last visit. The linear regression model was assumed for the observed data. In this thesis, we considered the PM-MI for incomplete longitudinal data and assumed the linear mixed effects model for the observed data. We have also carried out simulation studies to evaluate the performance of the sensitivity analyses approaches (Iddrisu and Gumedze, 2019*a*).

Finally, we proposed and conducted sensitivity analysis for the shared-parameter model framework via the global influence approach. We have also conducted simulation studies to evaluate the performance of the shared-parameter model (Iddrisu and Gumedze, 2019*b*).

7.5 Conclusions

The problem of missing data is ever-present in longitudinal studies designs. Missing data may bias statistical inferences and hence the need to accommodate missing data in the statistical analyses in other to obtain valid inferences. The best way to handle missing data is to avoid the occurrence of missing data (NRC, 2010; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013). However, the occurrence of missing data is often beyond the control of the researcher. In other to draw inference in the presence of missing data, we have conducted sensitivity analyses to investigate sensitivity of statistical inferences to alternative assumptions about the missing data mechanisms. These sensitivity analyses were applied to the incomplete longitudinal CD4 count measurements from the IMPI trial (Mayosi et al., 2012, 2014).

This study is the first study to address the problem of the missing data in the CD4 count measurements from the IMPI trial. There is a rich literature on how to draw inferences with incomplete data. However, the literature on how to assess sensitivity to various forms of assumptions about the nature of the dropout mechanism is relatively new (NRC, 2010). This is because assessment of sensitivity to various assumptions is an active area of research and thus more difficult to identify a general agreement about how sensitivity analyses should be conducted (NRC, 2010).

In this thesis, we have made an effort to conduct sensitivity analysis to address the problem of the incomplete CD4 count data from the IMPI trial. We have not derived new statistical theory. One of our main challenges was to cope with the unobserved CD4 counts measurements in the IMPI clinical trial. We had to find suitable sensitivity analysis approaches that contextually addresses the problem of the unobserved CD4 count measurements. We recognized that this case study cannot cover the broad range of types and designs of clinical trials because the literature on sensitivity analysis to nonrandom dropout is evolving. Hence, the primary objective of this thesis is to assert the importance of conducting some form of sensitivity analysis and to illustrate the principles in the IMPI trial setting.

We encourage researchers to carefully select primary analysis model and then consider methods that stress-test the primary analysis model's assumption via sensitivity analysis (Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013; Iddrisu and Gumedze, 2018). In this way, the researcher is able to reveal the results of sensitivity analysis to various deviations from assumptions made about the missing data mechanism. Such outcomes can be discussed with clinical reviewers to determine if they are implausibly unfavorable.

We have demonstrated that the statistical inferences, from the MAR primary analysis model, are robust to the NMAR assumptions and influential subjects do not overturn the study conclusions about treatment effects and the dropout mechanism. This means that even under NMAR assump-

tions, the conclusions are essentially unchanged and that the MAR analysis provides a scientifically useful statistical inferences whose conclusions are robust over contextually plausible assumptions about the missing data.

Apart from highlighting the relative importance of various tools available to the researcher, this thesis also emphasizes the practicality and importance of conducting sensitivity analyses. Various approaches are introduced, explored and illustrated especially in the face of uncertainty about the underlying missing data mechanism, the application of several approaches can only increase the researcher's confidence in the resulting conclusions, and is therefore strongly recommended.

The IMPI trial was a cardiology trial and HIV-related data were collected. However, the HIV data were not collected as would have been in a HIV focused clinical trial, and hence there are missing CD4 count. Despite the fact that the IMPI trial is a cardiology trial, our analyses of the HIV data provide reasonable information regarding the effect of prednisolone on CD4 count changes over time. The missingness of CD4 values might be informative, and hence later values of CD4 count might be missing because 246 (17.57%) of the 1400 patients in the IMPI trial died. This would require joint modelling (Baghfalaki et al., 2017; De Gruttola and Tu, 1994; Rizopoulos et al., 2008; Little, 1995) on the CD4 count data and time to death. The JM for longitudinal and survival outcomes is beyond the scope of this thesis.

There is no universal approach for conducting sensitivity analysis to missing data since the objective of clinical trial varies from trial to trial as well as the method of analysis (NRC, 2010). However, the National Research Council (NRC, 2010) panel provided a number of guidelines under which appropriate sensitivity analysis can conducted, given the trial setting. The guidelines and recommendations especially emphasise on the importance of carefully selecting primary analysis methods based on clearly formulated assumptions regarding the missingness mechanism and on the necessity to perform a range of sensitivity analyses that stress-test the results of the primary analysis under different sets of assumptions (Ratitch et al., 2013). Sensitivity analysis to missing data should be considered as default when analysing data with missing values (Thabane et al., 2013). So recognizing the importance and need for rigorous sensitivity analysis to missing data, we considered three modelling frameworks made up of different assumptions about the missing data. We used these frameworks to conduct sensitivity analysis to the missing CD4 count data in the IMPI trial (Mayosi et al., 2012, 2014; Iddrisu and Gumedze, 2018, 2019*a*,*b*).

A careful study of these frameworks for conducting sensitivity analysis as well as their applications in trial settings provides a tool for proper understanding on which sensitivity analysis method to use and in what circumstance. For instance, given a trial setting, one may consider sensitivity analysis that assumes that dropouts' responses, in a given treatment arm, are worsened by some amount (see Ratitch et al. (2013) paper). Usually, the assumption of worse responses (for dropouts) applies to the

active treatment arm(s), while responses for dropouts in the other treatment arms may be imputed based on MAR. Also, in a trial setting where dropouts receive treatment (standard of care) assumed to be equivalent to the randomized treatment, dropouts' responses should be adjusted comparable with those who stayed in the trial and received randomized treatment (Carpenter et al., 2013; Mallinckrodt, Roger, Chuang-Stein, Molenberghs, O'Kelly, Ratitch, Janssens and Bunouf, 2013). In this case, one can assumed that the response distribution of the dropouts is the same as those in the trial.

The third scenario of a trial setting is where the responses (of patients in both the treatment and placebo arms) under consideration are expected to be comparable (Iddrisu and Gumedze, 2019*a*; Carpenter et al., 2013). In this scenario, sensitivity analyses can be design to assume that placebo dropouts evolve in the same way as patients in the treatment arm that remain in the study and vice versa (see analyses in Iddrisu and Gumedze (2019*a*) paper and Chapter 5 of this thesis).

In a trial setting where patients are removed/unobserved from a study due to designed reason but not a genuine dropout (see milk protein trial in Verbyla and Cullis (1990)), sensitivity analyses (see Diggle and Kenward (1994); Kenward (1998); Verbeke et al. (2001); Shen et al. (2006); Thijs et al. (2000); Iddrisu and Gumedze (2018, 2019*b*) papers and Chapters 4 and 6 of this thesis) can be conducted to identify and investigate the influence of influential patients on statistical inferences. Analysts do not necessarily have to implement all the sensitivity analysis approaches explored in this thesis in order to draw statistical inferences, but have to clearly defined the trial setting (design) so that conclusion about which of the methods to use can be made.

7.6 Further research

The DK selection model methodology considered in this study is used to assess the effect of influential subjects on parameter estimates and conclusions (Thijs et al., 2000; Verbeke et al., 2001; Shen et al., 2006). However, the methodology can also be used to investigate the influence of individual's observations on model estimation.

In addition, there are other methods of identifying outliers/influential observations which could be used within the DK model. Examples of such methods are the variance shift outlier model (VSOM) (Gumedze et al., 2010) and Cook's distance (Cook, 1986).

The DK selection assumes that the probability of dropout depends on the response variable of interest. An alternative parametrization of the dropout probability in terms of residuals (rather than the response) would produce a different picture (Verbeke et al., 2001). However, this parametrization would require a new theoretical development (Verbeke and Molenberghs, 2009, Ch. 19, pp. 311-312).

References

Albert, P. S. and Follmann, D. A. (2000), Modeling repeated count data subject to informative dropout, *Biometrics* **56**(3), 667–677.

Albert, P. S., Follmann, D. A., Wang, S. A. and Suh, E. B. (2002), A latent autoregressive model for longitudinal binary data subject to informative missingness, *Biometrics* **58**(3), 631–642.

Alfò, M. and Aitkin, M. (2000), Random coefficient models for binary longitudinal responses with attrition, *Statistics and Computing* **10**(4), 279–287.

Ali, A., Dawson, S., Blows, F., Provenzano, E., Ellis, I., Baglietto, L., Huntsman, D., Caldas, C. and Pharoah, P. (2011), Comparison of methods for handling missing data on immunohistochemical markers in survival analysis of breast cancer, *British Journal of cancer* **104**(4), 693–699.

Ayele, B. T., Lipkovich, I., Molenberghs, G. and Mallinckrodt, C. H. (2014), A multiple-imputation-based approach to sensitivity analyses and effectiveness assessments in longitudinal clinical trials, *Journal of Biopharmaceutical Statistics* **24**(2), 211–228.

Baghfalaki, T., Ganjali, M. and Verbeke, G. (2017), A shared parameter model of longitudinal measurements and survival time with heterogeneous random-effects distribution, *Journal of Applied Statistics* **44**(15), 2813–2836.

Bahadur, R. R. (1961), A representation of the joint distribution of responses to n dichotomous items, *Studies in item analysis and prediction* **6**, 158–168.

Beckman, R. J., Nachtsheim, C. J. and Cook, R. D. (1987), Diagnostics for mixed–model analysis of variance, *Technometrics* **29**(4), 413–426.

Bernhard, J., Cella, D. F., Coates, A. S., Fallowfield, L., Ganz, P. A., Moinpour, C. M., Mosconi, P., Osoba, D., Simes, J. and Hürny, C. (1998), Missing quality of life data in cancer clinical trials: serious problems and challenges, *Statistics in medicine* **17**(5-7), 517–532.

Beunckens, C., Molenberghs, G., Verbeke, G. and Mallinckrodt, C. (2008), A latent-class mixture model for incomplete longitudinal gaussian data, *Biometrics* **64**(1), 96–105.

Box, G. E. (1950), Problems in the analysis of growth and wear curves, *Biometrics* **6**(4), 362–389.

Carpenter, J. and Goldstein, H. (2004), Multiple imputation in mlwin, *Multilevel modelling newsletter* **16**(2), 9–18.

Carpenter, J. and Kenward, M. (2012), *Multiple imputation and its application*, John Wiley and Sons.

Carpenter, J., Kenward, M., Evans, S. and White, I. (2003), Letter to the editor: Last observation carry forward and last observation analysis, *Statistics in Medicine* **22**, 2429–2441.

Carpenter, J. R. and Kenward, M. G. (2007), Missing data in randomised controlled trials-a practical guide, *London School of Hygiene* .

Carpenter, J. R., Kenward, M. G. and Vansteelandt, S. (2006), A comparison of multiple imputation and doubly robust estimation for analyses with missing data, *Journal of the Royal Statistical Society: Series A* **169**(3), 571–584.

Carpenter, J. R., Kenward, M. G. and White, I. R. (2007), Sensitivity analysis after multiple imputation under missing at random: a weighting approach, *Statistical Methods in Medical Research* **16**(3), 259–275.

Carpenter, J. R., Roger, J. H. and Kenward, M. G. (2013), Analysis of longitudinal trials with protocol deviation: A framework for relevant, accessible assumptions, and inference via multiple imputation, *Journal of Biopharmaceutical Statistics* **23**(6), 1352–1371.

Chatterjee, S. and Hadi, A. S. (2009), *Sensitivity analysis in linear regression*, John Wiley and Sons.

CHMP (2010), Committee for medicinal products for human use (chmp), guideline on missing, data in confirmatory clinical trials, *European Medicines Agency,London, UK* .

Chow, S.-C. and Liu, J.-P. (2008), *Design and analysis of clinical trials: concepts and methodologies*, John Wiley and Sons.

Cole, B. F., Bonetti, M., Zaslavsky, A. M. and Gelber, R. D. (2005), A multistate markov chain model for longitudinal, categorical quality-of-life data subject to non-ignorable missingness, *Statistics in Medicine* **24**(15), 2317–2334.

Cook, R. D. (1986), Assessment of local influence, *Journal of the Royal Statistical Society. Series B* **48**(2), 133–169.

Cook, R. D. and Weisberg, S. (1982), Residuals and influence in regression, *New York: Chapman and Hall* .

Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and Kenward, M. G. (2010), Generalized shared-parameter models and missingness at random, *Statistical modelling* **11**(4), 279–310.

Creemers, A., Hens, N., Aerts, M., Molenberghs, G., Verbeke, G. and Kenward, M. G. (2011), A sensitivity analysis for shared-parameter models for incomplete longitudinal outcomes, *Biometrical Journal* **52**(1), 111–125.

Daniels, M. J. and Hogan, J. W. (2008), *Missing data in longitudinal studies: Strategies for Bayesian modeling and sensitivity analysis*, CRC Press.

De Gruttola, V. and Tu, X. M. (1994), Modelling progression of cd4-lymphocyte count and its relationship to survival time, *Biometrics* **50**(4), 1003–1014.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), Maximum likelihood from incomplete data via the em algorithm, *Journal of the Royal Statistical Society. Series B* **39**(1), 1–38.

Diggle, P., Heagerty, P., Liang, K.-Y. and Zeger, S. (2002), *Analysis of longitudinal data*, Oxford University Press.

Diggle, P. and Kenward, M. G. (1994), Informative drop-out in longitudinal data analysis, *Applied statistics* **43**(1), 49–93.

Fitzmaurice, G., Davidian, M., Verbeke, G. and Molenberghs, G. (2008), *Longitudinal data analysis*, CRC Press.

Fitzmaurice, G. M. and Laird, N. M. (2000), Generalized linear mixture models for handling nonignorable dropouts in longitudinal studies, *Biostatistics* **1**(2), 141–156.

Follmann, D. and Wu, M. (1995*a*), An approximate generalized linear model with random effects for informative missing data, *Biometrics* **51**, 151–168.

Follmann, D. and Wu, M. (1995*b*), An approximate generalized linear model with random effects for informative missing data, *Biometrics* **51**(1), 151–168.

Gao, S. (2004), A shared random effect parameter approach for longitudinal dementia data with non-ignorable missing data, *Statistics in Medicine* **23**(2), 211–219.

Glonek, G. (1999), On identifiability in models for incomplete binary data, *Statistics and Probability Letters* **41**(2), 191–197.

Greenland, S. and Finkle, W. D. (1995), A critical look at methods for handling missing covariates in epidemiologic regression analyses, *American Journal of Epidemiology* **142**(12), 1255–1264.

Grizzle, J. E. and Allen, D. M. (1969), Analysis of growth and dose response curves, *Biometrics* **25**(2), 357–381.

Gumedze, F. and Dunne, T. (2011), Parameter estimation and inference in the linear mixed model, *Linear Algebra and its Applications* **435**(8), 1920–1944.

Gumedze, F. N. (2008), A variance shilf model for outlier detection and estimation in linear and linear mixed models.

Gumedze, F. N., Welham, S. J., Gogel, B. J. and Thompson, R. (2010), A variance shift model for detection of outliers in the linear mixed model, *Computational Statistics and Data Analysis* **54**(9), 2128–2144.

Harville, D. A. (1974), Bayesian inference for variance components using only error contrasts, *Biometrika* **61**(2), 383–385.

Harville, D. A. (1977), Maximum likelihood approaches to variance component estimation and to related problems, *Journal of the American Statistical Association* **72**(358), 320–338.

Heckman, J. J. (1976), The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, **5**(4), 475–492.

Henderson, C. R. (1953), Estimation of variance and covariance components, *Biometrics* **9**(2), 226–252.

Henderson, C. R., Kempthorne, O., Searle, S. R. and Von Krosigk, C. (1959), The estimation of environmental and genetic trends from records subject to culling, *Biometrics* **15**(2), 192–218.

Hogan, J. W. and Laird, N. M. (1997), Mixture models for the joint distribution of repeated measures and event times, *Statistics in Medicine* **16**(3), 239–257.

Hughes, R., Sterne, J. and Tilling, K. (2016), Comparison of imputation variance estimators, *Statistical methods in medical research* **25**(6), 2541–2557.

Ibrahim, J. G. and Molenberghs, G. (2009), Missing data methods in longitudinal studies: a review, *Test* **18**(1), 1–43.

Iddrisu, A.-K. and Gumedze, F. (2018), Application of sensitivity analysis to incomplete longitudinal cd4 count data, *Journal of Applied Statistics* pp. 1–16.

Iddrisu, A.-K. and Gumedze, F. (2019*a*), An application of a pattern-mixture model with multiple imputation for the analysis of longitudinal trials with protocol deviations, *BMC medical research methodology* **19**(1), 10.

Iddrisu, A.-K. and Gumedze, F. (2019*b*), Sensitivity analysis for the generalized shared-parameter model framework, *Journal of Biopharmaceutical Statistics* pp. 1–19.

Jansen, I., Hens, N., Molenberghs, G., Aerts, M., Verbeke, G. and Kenward, M. G. (2006), The nature of sensitivity in monotone missing not at random models, *Computational Statistics and Data Analysis* **50**(3), 830–858.

Jansen, I., Molenberghs, G., Aerts, M., Thijs, H. and Van Steen, K. (2003), A local influence approach applied to binary data from a psychiatric study, *Biometrics* **59**(2), 410–419.

Jennrich, R. I. and Schluchter, M. D. (1986), Unbalanced repeated-measures models with structured covariance matrices, *Biometrics* **42**(4), 805–820.

Kenward, M. G. (1998), Selection models for repeated measurements with non-random dropout: an illustration of sensitivity, *Statistics in Medicine* **17**(23), 2723–2732.

Kidson, J. W. and Trenberth, K. E. (1988), Effects of missing data on estimates of monthly mean general circulation statistics, *Journal of Climate* **1**(12), 1261–1275.

Kim, J.-O. and Curry, J. (1977), The treatment of missing data in multivariate analysis, *Sociological Methods and Research* **6**(2), 215–240.

Laird, N. M. (1988), Missing data in longitudinal studies, *Statistics in Msedicine* **7**(1-2), 305–315.

Laird, N. M. and Ware, J. H. (1982), Random-effects models for longitudinal data, *Biometrics* **38**(4), 963–974.

Lavori, P. W., Dawson, R. and Shera, D. (1995), A multiple imputation strategy for clinical trials with truncation of patient data, *Statistics in Medicine* **14**(17), 1913–1925.

Lesaffre, E. and Verbeke, G. (1998), Local influence in linear mixed models, *Biometrics* **54**(2), 570–582.

Liang, K.-Y. and Zeger, S. L. (1986), Longitudinal data analysis using generalized linear models, *Biometrika* **73**(1), 13–22.

Lindstrom, M. J. and Bates, D. M. (1988), Newton raphson and em algorithms for linear mixed-effects models for repeated-measures data, *Journal of the American Statistical Association* **83**(404), 1014–1022.

Little, R. J. (1992), Regression with missing x's: a review, *Journal of the American Statistical Association* **87**(420), 1227–1237.

Little, R. J. (1993), Pattern-mixture models for multivariate incomplete data, *Journal of the American Statistical Association* **88**(421), 125–134.

Little, R. J. (1995), Modeling the drop-out mechanism in repeated-measures studies, *Journal of the American Statistical Association* **90**(431), 1112–1121.

Little, R. J. and Rubin, D. B. (1987), *Statistical analysis with missing data*, John Wiley ands Sons.

Little, R. and Yau, L. (1996), Intent-to-treat analysis for longitudinal studies with drop-outs, *Biometrics* **52**(4), 1324–1333.

Mallinckrod, C. H., Lane, P. W., Schnell, D., Peng, Y. and Mancuso, J. P. (2008), Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials, *Drug Information Journal* **42**(4), 303–319.

Mallinckrodt, C., Lin, Q. and Molenberghs, M. (2013), A structured framework for assessing sensitivity to missing data assumptions in longitudinal clinical trials, *Pharmaceutical Statistics* **12**(1), 1–6.

Mallinckrodt, C., Roger, J., Chuang-Stein, C., Molenberghs, G., O'Kelly, M., Ratitch, B., Janssens, M. and Bunouf, P. (2013), Recent developments in the prevention and treatment of missing data, *Therapeutic Innovation ands Regulatory Science* .

Mayosi, B. M., Ntsekhe, M., Bosch, J., Pandie, S., Jung, H., Gumedze, F., Pogue, J., Thabane, L., Smieja, M., Francis, V. et al. (2014), Prednisolone and mycobacterium indicus pranii in tuberculous pericarditis, *New England Journal of Medicine* **371**(12), 1121–1130.

Mayosi, B. M., Ntsekhe, M., Bosch, J., Pogue, J., Gumedze, F., Badri, M., Jung, H., Pandie, S., Smieja, M., Thabane, L. et al. (2012), Rationale and design of the investigation of the management of pericarditis (impi) trial: A 2× 2 factorial randomized double-blind multicenter trial of adjunctive prednisolone and mycobacterium w immunotherapy in tuberculous pericarditis, *American Heart Journal* **165**, 109–115.

Michiels, B., Molenberghs, G., Bijnens, L., Vangeneugden, T. and Thijs, H. (2002), Selection models and pattern-mixture models to analyse longitudinal quality of life data subject to drop-out, *Statistics in Medicine* **21**(8), 1023–1041.

Miller, J. J. (1977), Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance, *The Annals of Statistics* **5**(4), 746–762.

Molenberghs, G., Beunckens, C., Sotto, C. and Kenward, M. G. (2008), Every missingness not at random model has a missingness at random counterpart with equal fit, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **70**(2), 371–388.

Molenberghs, G., Fitzmaurice, G., Kenward, M. G., Tsiatis, A. and Verbeke, G. (2014), *Handbook of Missing Data Methodology*, CRC Press.

Molenberghs, G. and Kenward, M. (2007), *Missing data in clinical studies*, John Wiley & Sons.

Molenberghs, G., Kenward, M. G. and Goetghebeur, E. (2001), Sensitivity analysis for incomplete contingency tables: the slovenian plebiscite case, *Journal of the Royal Statistical Society: Series C* **50**(1), 15–29.

Molenberghs, G., Kenward, M. G. and Lesaffre, E. (1997), The analysis of longitudinal ordinal data with nonrandom drop-out, *Biometrika* **84**(1), 33–44.

Molenberghs, G. and Lesaffre, E. (1994), Marginal modeling of correlated ordinal data using a multivariate plackett distribution, *Journal of the American Statistical Association* **89**(426), 633–644.

Molenberghs, G., Michiels, B., Kenward, M. G. and Diggle, P. J. (1998), Monotone missing data and pattern-mixture models, *Statistica Neerlandica* **52**(2), 153–161.

Molenberghs, G., Thijs, H., Jansen, I., Beunckens, C., Kenward, M. G., Mallinckrodt, C. and Carroll, R. J. (2004), Analyzing incomplete longitudinal clinical trial data, *Biostatistics* **5**(3), 445–464.

Molenberghs, G. and Verbeke, G. (2006), *Models for discrete longitudinal data*, Springer Science and Business Media.

Molenberghs, G., Verbeke, G., Thijs, H., Lesaffre, E. and Kenward, M. G. (2001), Influence analysis to assess sensitivity of the dropout process, *Computational Statistics and Data Analysis* **37**(1), 93–113.

Mori, M., Woolson, R. F. and Woodworth, G. G. (1994), Slope estimation in the presence of informative right censoring: modeling the number of observations as a geometric random variable, *Biometrics* **50**(1), 39–50.

Morris, T. P., White, I. R. and Crowther, M. J. (2019), Using simulation studies to evaluate statistical methods, *Statistics in medicine* **38**(11), 2074–2102.

Morris, T. P., White, I. R. and Royston, P. (2014), Tuning multiple imputation by predictive mean matching and local residual draws, *BMC medical research methodology* **14**(1), 75.

NRC (2010), The prevention and treatment of missing data in clinical trials. panel on handling missing data in clinical trials. committee on national statistics, division of behavioral and social sciences and education, *The National Academies Press: Washington DC* .

O'neill, B. (2006), *Elementary differential geometry*, Academic press.

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

Permutt, T. (2015), Sensitivity analysis for missing data in regulatory submissions, *Statistics in medicine* .

Permutt, T. (2016), Sensitivity analysis for missing data in regulatory submissions, *Statistics in medicine* **35**(17), 2876–2879.

Potthoff, R. F. and Roy, S. (1964), A generalized multivariate analysis of variance model useful especially for growth curve problems, *Biometrika* **51**(3-4), 313–326.

Pulkstenis, E. P., Ten Have, T. R. and 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.

Rakhmawati, T. W., Molenberghs, G., Verbeke, G. and Faes, C. (2016), Local influence diagnostics for incomplete overdispersed longitudinal counts, *Journal of Applied Statistics* **43**(9), 1722–1737.

Rakhmawati, T. W., Molenberghs, G., Verbeke, G. and Faes, C. (2017), Local influence diagnostics for generalized linear mixed models with overdispersion, *Journal of Applied Statistics* **44**(4), 620–641.

Rao, C. R. (1958), Some statistical methods for comparison of growth curves, *Biometrics* **14**(1), 1–17.

Rao, C. R. (1965), The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, *Biometrika* **52**(4), 447–458.

Ratitch, B., O'Kelly, M. and Tosiello, R. (2013), Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models, *Pharmaceutical Statistics* **12**(6), 337–347.

Rezvan, P. H., White, I. R., Lee, K. J., Carlin, J. B. and Simpson, J. A. (2015), Evaluation of a weighting approach for performing sensitivity analysis after multiple imputation, *BMC medical research methodology* **15**(1), 83.

Rizopoulos, D., Verbeke, G. and Molenberghs, G. (2008), Shared parameter models under random effects misspecification, *Biometrika* **95**(1), 63–74.

Robins, J. M., Greenland, S. and Hu, F.-C. (1999), Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome, *Journal of the American Statistical Association* **94**(447), 687–700.

Robins, J. M. and Rotnitzky, A. (1995), Semiparametric efficiency in multivariate regression models with missing data, *Journal of the American Statistical Association* **90**(429), 122–129.

Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1995), Analysis of semiparametric regression models for repeated outcomes in the presence of missing data, *Journal of the American Statistical Association* **90**(429), 106–121.

Roy, J. (2003), Modeling longitudinal data with nonignorable dropouts using a latent dropout class model, *Biometrics* **59**(4), 829–836.

Royston, P. (2005), Multiple imputation of missing values: update of ice, *Stata Journal* **5**(4), 527.

Rubin, D. B. (1976), Inference and missing data, *Biometrika* **63**(3), 581–592.

Rubin, D. B. (1977), Formalizing subjective notions about the effect of nonrespondents in sample surveys, *Journal of the American Statistical Association* **72**(359), 538–543.

Rubin, D. B. (1987), The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The sir algorithm, *Journal of the American Statistical Association* **82**(398), 543–546.

Rubin, D. B. (1996), Multiple imputation after 18+ years, *Journal of the American Statistical Association* **91**(434), 473–489.

Schafer, J. L. and Graham, J. W. (2002), Missing data: our view of the state of the art., *Psychological methods* **7**(2), 147.

Scharfstein, D. O., Rotnitzky, A. and Robins, J. M. (1999), Adjusting for nonignorable dropout using semiparametric nonresponse models, *Journal of the American Statistical Association* **94**(448), 1096–1120.

Scheffe, H. (1999), *The analysis of variance*, Vol. 72, John Wiley and Sons.

Schluchter, M. D. (1992), Methods for the analysis of informatively censored longitudinal data, *Statistics in Medicine* **11**(14-15), 1861–1870.

Seaman, S. R. and White, I. R. (2013), Review of inverse probability weighting for dealing with missing data, *Statistical Methods in Medical Research* **22**(3), 278–295.

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

Searle, S. R. (2012), *Linear models*, Wiley.

Searle, S. R., Casella, G. and McCulloch, C. E. (2009), *Variance components*, Vol. 391, John Wiley & Sons.

Shen, S., Beunckens, C., Mallinckrodt, C. and Molenberghs, G. (2006), A local influence sensitivity analysis for incomplete longitudinal depression data, *Journal of Biopharmaceutical Statistics* **16**(3), 365–384.

Shi, X., Zhu, H. and Ibrahim, J. G. (2009), Local influence for generalized linear models with missing covariates, *Biometrics* **65**(4), 1164–1174.

Smuk, M. (2015), Missing data methodology: sensitivity analysis after multiple imputation, PhD thesis, Ph. D. Thesis.

Sotto, C. (2009), Topics in analysis and sensitivity analysis of incomplete longitudinal data.

Su, Y.-S., Yajima, M., Gelman, A. E. and Hill, J. (2011), Multiple imputation with diagnostics (mi) in r: Opening windows into the black box, *Journal of Statistical Software* **45**(2), 1–31.

Thabane, L., Mbuagbaw, L., Zhang, S., Samaan, Z., Marcucci, M., Ye, C., Thabane, M., Giangregorio, L., Dennis, B., Kosa, D. et al. (2013), A tutorial on sensitivity analyses in clinical trials: the what, why, when and how, *BMC Medical Research Methodology* **13**(1), 92.

Thijs, H., Molenberghs, G., Michiels, B., Verbeke, G. and Curran, D. (2002), Strategies to fit pattern-mixture models, *Biostatistics* **3**(2), 245–265.

Thijs, H., Molenberghs, G. and Verbeke, G. (2000), The milk protein trial: influence analysis of the dropout process, *Biometrical Journal* **42**(5), 617–646.

Thisted, R. A. (1988), *Elements of statistical computing: numerical computation*, CRC Press.

Thomas, R., Have, T., Kunselman, A. R., Pulkstenis, E. P. and Landis, J. R. (1998), Mixed effects logistic regression models for longitudinal binary response data with informative drop-out, *Biometrics* **54**(1), 367–383.

Troxel, A. B., Ma, G. and Heitjan, D. F. (2004), An index of local sensitivity to nonignorability, *Statistica Sinica* **14**(4), 1221–1238.

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

Tsiatis, A., Degruttola, V. and Wulfsohn, M. (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.

Tsonaka, R., Verbeke, G. and Lesaffre, E. (2009), A semi-parametric shared parameter model to handle nonmonotone nonignorable missingness, *Biometrics* **65**(1), 81–87.

Van Buuren, S. (2007), Multiple imputation of discrete and continuous data by fully conditional specification, *Statistical Methods in Medical Research* **16**(3), 219–242.

Van Buuren, S. (2012), *Flexible imputation of missing data*, CRC press.

Van Steen, K., Molenberghs, G., Verbeke, G. and Thijs, H. (2001), A local influence approach to sensitivity analysis of incomplete longitudinal ordinal data, *Statistical Modelling* **1**(2), 125–142.

Verbeke, G. and Molenberghs, G. (2009), *Linear mixed models for longitudinal data*, Springer Science and Business Media.

Verbeke, G. and Molenberghs, G. (2012), *Linear mixed models in practice: a SAS-oriented approach*, Springer Science and Business Media.

Verbeke, G., Molenberghs, G., Thijs, H., Lesaffre, E. and Kenward, M. G. (2001), Sensitivity analysis for nonrandom dropout: a local influence approach, *Biometrics* **57**(1), 7–14.

Verbyla, A. and Cullis, B. R. (1990), Modelling in repeated measures experiments, *Applied Statistics* **39**(3), 341–356.

Viviani, S., Rizopoulos, D. and Alfó, M. (2014), Local sensitivity to non-ignorability in joint models, *Statistical Modelling* **14**(3), 205–228.

Vonesh, E. F., Wang, H., Nie, L. and Majumdar, D. (2002), Conditional second-order generalized estimating equations for generalized linear and nonlinear mixed-effects models, *Journal of the American Statistical Association* **97**(457).

Wang-Clow, F., Lange, M., Laird, N. M. and Ware, J. H. (1995), A simulation study of estimators for rates of change in longitudinal studies with attrition, *Statistics in Medicine* **14**(3), 283–297.

Waternaux, C., Laird, N. M. and Ware, J. H. (1989), Methods for analysis of longitudinal data: blood-lead concentrations and cognitive development, *Journal of the American Statistical Association* **84**(405), 33–41.

West, B. T., Welch, K. B. and Galecki, A. T. (2014), *Linear mixed models: a practical guide using statistical software*, CRC Press.

White, I. (1997), Letter to the editor: Survival analysis of randomized clinical trials adjusted for patients who switch treatments by mg law and jm kaldor, statistics in medicine, 15, 2069–2076 (1996), *Statistics in medicine* **16**(22), 2619–2620.

White, I. R. and Thompson, S. G. (2005), Adjusting for partially missing baseline measurements in randomized trials, *Statistics in medicine* **24**(7), 993–1007.

Williamson, J. (2006), Models for discrete longitudinal data, *Journal of the American Statistical Association* **101**(475), 1307–1307.

Wishart, J. (1938), Growth-rate determinations in nutrition studies with the bacon pig, and their analysis, *Biometrika* **30**(2), 16–28.

Wu, M. C. and Bailey, K. (1988), Analysing changes in the presence of informative right censoring caused by death and withdrawal, *Statistics in Medicine* **7**(1-2), 337–346.

Wu, M. C. and Bailey, K. R. (1989), Estimation and comparison of changes in the presence of informative right censoring: conditional linear model, *Biometrics* **45**(3), 939–955.

Wu, M. C. and Carroll, R. J. (1988), Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process, *Biometrics* **44**(1), 175–188.

Wu, M. C. and Follmann, D. A. (1999), Use of summary measures to adjust for informative missingness in repeated measures data with random effects, *Biometrics* **55**(1), 75–84.

Xie, H., Gao, W., Xing, B., Heitjan, D. F., Hedeker, D. and Yuan, C. (2018), Measuring the impact of nonignorable missingness using the r package isni, *Computer methods and programs in biomedicine* **164**, 207–220.

Yuan, Y. and Little, R. J. (2009), Mixed-effect hybrid models for longitudinal data with nonignorable dropout, *Biometrics* **65**(2), 478–486.

Zhu, X. (2014), Comparison of four methods for handing missing data in longitudinal data analysis through a simulation study, *Open Journal of Statistics* **4**(11), 933.

Appendix A

SAS code for implementing sensitivity analyses for the DK selection model

This section presents SAS code and description of the code implemented to the sensitivity analyses under the DK selection model.

SAS implementation

The SAS code was implemented to CD4 count measurement data (IMPI_29_2016.sas) from the IMPI trial.

Data structure

The longitudinal data should be in a long format:

```
Data IMPI_2016;
input id sqcd4 transage art pred time;
CARDS;
1 15.26434 33 0 1 0
1 11.7047  33 0 1 1
1 15.52417 33 0 1 0.5
1 16.8226  33 1 1 3
1 18.02776 33 1 1 6
2 15.32971 32 0 1 0
2 14.79865 32 0 1 1
2 15.32971 32 0 1 0.5
2 21.56386 32 1 1 3
2 25.17936 32 1 1 6
. . . . . .

RUN;
```

Data variables description

Id: subject identification number.

sqcd4: square root of CD4 count.

transage: age of a subject.

art: anti-retroviral therapy.

pred: prednisolone

time: other forms of variables can be added depending on the model to be fitted.

Create the design matrix for the covariates $\mathbf{X}_i$ and a vector for the response variable $\mathbf{Y}_i$:

```
Proc iml;
use IMPI_2016;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5]; /*prednisolone*/
timetreat =time#treat; /*prednisolone x time interaction*/
art=data[,4];
predart= treat#art; /*prednisolone x time interaction*/
transage = data[,3];
intercept=j(nrow(data),1,1);

create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
```

Now supply the number of subjects (nsub), number of time points (ntime), and initial values for parameter estimates and variance-covariance structure; eg,

```
/*initial values for parameter estimates*/
beta= 24.3408//0.4241//0.2943//0.02986//1.9891//-0.05294//-3.2065;
/*variance-covariance matrix*/
sigma11=27.5557;
sigma21=23.1566;
sigma22=29.7099;
sigma31=21.7018;
sigma32=22.8297;
sigma33=29.6190;
sigma41=17.7755;
sigma42=17.6218;
sigma43=20.0637;
sigma44=25.7221;
sigma51 = 18.4084;
sigma52 = 18.7202;
sigma53 = 20.3257;
sigma54 = 22.7242;
sigma55 = 26.2859;
v = j(5,5,1);
v[1,1] = sigma11;
v[1,2] = sigma21;
v[2,1] = sigma21;
v[2,2] = sigma22;
v[1,3] = sigma31;
v[3,1] = sigma31;
v[2,3] = sigma32;
v[3,2] = sigma32;
v[3,3] = sigma33;
v[1,4] = sigma41;
v[4,1] = sigma41;
v[2,4] = sigma42;
v[4,2] = sigma42;
v[3,4] = sigma43;
v[4,3] = sigma43;
v[4,4] = sigma44;
v[5,1] = sigma51;
v[1,5] = sigma51;
v[5,2] = sigma52;
v[2,5] = sigma52;
v[5,3] = sigma53;
v[3,5] = sigma53;
v[5,4] = sigma54;
v[4,5] = sigma54;
v[5,5] = sigma55;
u=root(v);
v2=t(u)*u;
uvec=colvec(u);
```

```
nozero=loc(uvec<>0);
uvec=uvec[nozero];
psi= 1; /*initial value for intercept in the dropout model*/
initial=beta//uvec//psi;
create initial var {initial};
append;
nsub=137;/*number of subjects*/
ntime=5; /*number of time points*/
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;
```

For ψ_0, ψ_1, and ψ_2, initial values are provided in a vector form as $[1:3]$. This means that under MCAR, only ψ_0 is required and ψ_0 and ψ_2 are required for MAR assumption, and ψ_0, ψ_1, and ψ_2 are required for NMAR mechanism. The above information is then used to fit the DK selection the local influence models under MCAR, MAR, and NMAR mechanisms.

```
/************************************/
/* (1) DIGGLE-KENWARD MODEL : MCAR   */
/************************************/

proc iml;
use IMPI_2016;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
beta= 24.3408//0.4241//0.2943//0.02986//1.9891//-0.05294//-3.2065;
sigma11=27.5557;
sigma21=23.1566;
sigma22=29.7099;
sigma31=21.7018;
sigma32=22.8297;
sigma33=29.6190;
sigma41=17.7755;
sigma42=17.6218;
sigma43=20.0637;
sigma44=25.7221;
sigma51 = 18.4084;
sigma52 = 18.7202;
sigma53 = 20.3257;
sigma54 = 22.7242;
sigma55 = 26.2859;
v = j(5,5,1);
v[1,1] = sigma11;
v[1,2] = sigma21;
v[2,1] = sigma21;
v[2,2] = sigma22;
v[1,3] = sigma31;
v[3,1] = sigma31;
v[2,3] = sigma32;
v[3,2] = sigma32;
v[3,3] = sigma33;
v[1,4] = sigma41;
v[4,1] = sigma41;
v[2,4] = sigma42;
v[4,2] = sigma42;
v[3,4] = sigma43;
v[4,3] = sigma43;
v[4,4] = sigma44;
v[5,1] = sigma51;
```

```
v[1,5] = sigma51;
v[5,2] = sigma52;
v[2,5] = sigma52;
v[5,3] = sigma53;
v[3,5] = sigma53;
v[5,4] = sigma54;
v[4,5] = sigma54;
v[5,5] = sigma55;
u=root(v);
v2=t(u)*u;
uvec=colvec(u);
nozero=loc(uvec<>0);
uvec=uvec[nozero];
psi= 1;
initial=beta//uvec//psi;
create initial var {initial};
append;
nsub=137;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;



proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1]=parameters[23];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */
```

```
xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);

/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */


if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
```

```
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);


ll=ll+ll1+ll2+ll3;
```

```
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,23,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;

data impi2aMCARcovar;
set covmatrix;
run;
proc print data=impi2aMCARcovar;
run;

data impi2aMCAR;
set result;
run;
proc print data=impi2aMCAR;
run;

data impi2aMCARhessian;
set hessianmatrix;
run;
proc print data=impi2aMCARhessian;
run;
data impi2aMCARloglik;
set loglik;
run;
proc print data=impi2aMCARloglik;
run;

/****************************************/
/* (2)  DIGGLE-KENWARD MODEL UNDER MAR   */
/****************************************/
proc iml;
use IMPI_2016;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
```

```
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMCAR;
read all var{est} into mcarest;
*initial=mcarest//0.5;
initial=mcarest//0.05;
create initial var {initial};
append;
nsub=137;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:2]=parameters[23:24];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
```

```
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);

/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
```

```
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
```

```
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,24,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
data impi2aMARcovar;
set covmatrix;
run;
proc print data=impi2aMARcovar;
run;

data impi2aMAR;
set result;
run;
proc print data=impi2aMAR;
run;

data impi2aMARhessian;
set hessianmatrix;
run;
proc print data=impi2aMARhessian;
run;
data impi2aMARloglik;
set loglik;
run;
proc print data=impi2aMARloglik;
run;


/************************************/
/* (3) DIGGLE-KENWARD MODEL : MNAR   */
/************************************/
proc iml;
use IMPI_2016;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
```

```
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMAR;
read all var{est} into mcarest;
*initial=mcarest//0.5;
initial=mcarest//0.05;
create initial var {initial};
append;
nsub=137;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;
proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:3]=parameters[23:25];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
```

```
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);

/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
```

```
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);


ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
```

```
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,25,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
data impi2aNMARcovar;
set covmatrix;
run;
proc print data=impi2aNMARcovar;
run;

data impi2aNMAR;
set result;
run;
proc print data=impi2aNMAR;
run;

data impi2aNMARhessian;
set hessianmatrix;
run;
proc print data=impi2aNMARhessian;
run;
data impi2aNMARloglik;
set loglik;
run;
proc print data=impi2aNMARloglik;
run;



/*%%%%%%%%%%%%%%%%%%%%%%%/
 (4) LOCAL INFLUENCE
/*%%%%%%%%%%%%%%%%%%%%%%%*/

proc iml;
use IMPI_29_2016;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
```

```
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMCAR;
read all var{est} into mcarest;
*initial=mcarest//0.5;
initial=mcarest//0.05;
create initial var {initial};
append;
nsub=137;
ntime=4;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u22=parameters[12];
u23=parameters[13];
u24=parameters[14];
u33=parameters[15];
u34=parameters[16];
u44=parameters[17];
psi=j(3,1,0);
psi[1:2]=parameters[18:19];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[2,1] = 0;
ui[2,2] = u22;
```

```
ui[2,3] = u23;
ui[2,4] = u24;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);

/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);
```

```
/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;


start delta(parameters) global(lastobs,vcurr,ecurr,x,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u22=parameters[12];
```

```
u23=parameters[13];
u24=parameters[14];
u33=parameters[15];
u34=parameters[16];
u44=parameters[17];
psi[1:2]=parameters[18:19]; nbeta=nrow(beta);
nsigma=10;
npsi=2;
ind=1;
ll=0;
ind=1;

do while (ind<=nsub);
/* select information on one particular patient */
xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
vi=t(ui)*ui;
sigma1=vi[1,1];
sigma12=vi[1,2];
sigma13=vi[1,3];
sigma14=vi[1,4];
sigma2=vi[2,2];
sigma23=vi[2,3];
sigma24=vi[2,4];
sigma3=vi[3,3];
sigma34=vi[3,4];
sigma4=vi[4,4];
psi=j(3,1,0);
pi = constant('PI');
g=j(nsub,1,0);

/* yobs and dropout */

if (yi[nrow(yi)]=.) then do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else do;
dropout=0;
yobs=yi;
end;

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
```

```
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* calculation of delta  */

contribd = j(1,npsi,0);
contribm = j(nbeta+nsigma,1,0);

if dropout=0 then do;  /* COMPLETER */
j=2;
do while (j <= nrow(yobs));
g = exp(psi[1]+yobs[j-1]*psi[2]);
g = g/(g+1);
hij = 1||yobs[j-1];
contribd = contribd-yobs[j]#hij#g#(1-g);
j=j+1;
end;
end;
else do;         /* INCOMPLETER */
if (di=2) then do;
hij=yobs;
g = exp(psi[1]+yobs[nrow(yobs)]*psi[2]);
g = g/(g+1);
muc = mui[di];
mup = mui[nummi];
hist = yobs;
vi11 = vi[nummi,nummi];
vi21 = vi[di,nummi];
vi11i = inv(vi11);
ecurr = muc + vi21*(vi11i)*(hist-mup);
contribd = contribd - ecurr#hij#g#(1-g);
/* use derivative wrt beta */
contribm[1:nbeta] = (1-g)#(xi[di,]-vi21*vi11i*xi[nummi,])';
/* use derivative wrt sigma1 */
contribm[nbeta+1] = (1-g)#(0-vi21*vi11i*2*sigma1)*vi11i*(hist-mup);
/* use derivative wrt sigma12 */
contribm[nbeta+2] = (1-g)#(1-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma13 */
contribm[nbeta+3] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma14 */
contribm[nbeta+4] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma2 */
contribm[nbeta+5] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma23 */
contribm[nbeta+6] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma24 */
contribm[nbeta+7] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma3 */
contribm[nbeta+8] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma34 */
contribm[nbeta+9] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma4 */
contribm[nbeta+10] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
end;
else if (di=3) then do;
j=2;
do while (j <= nrow(yobs));
g = exp(psi[1]+yobs[j-1]*psi[2]);
g = g/(g+1);
hij = 1||yobs[j-1];
contribd = contribd-yobs[j]#hij#g#(1-g);
j = j+1;
end;
g = exp(psi[1]+yobs[nrow(yobs)]*psi[2]);
g = g/(g+1);
muc = mui[di];
mup = mui[nummi];
hist = yobs;
vi11 = vi[nummi,nummi];
vi21 = vi[di,nummi];
vi11i = inv(vi11);
ecurr = muc + vi21*(vi11i)*(hist-mup);
```

```
contribd = contribd - ecurr#hij#g#(1-g);
/* use derivative wrt beta */
contribm[1:nbeta] = (1-g)#(xi[di,]-vi21*vi11i*xi[nummi,])';
/* use derivative wrt sigma1 */
vi21d=0||0;
vi21d = vi21d[nummi]';
helpvector=2*sigma1||0||0||0;
vi11d=shape(helpvector,2,2);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+1] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma12 */
vi21d=0||0;
vi21d = vi21d[nummi]';
helpvector=0||1||1||0;
vi11d=shape(helpvector,2,2);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+2] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma13 */
vi21d=1||0;
vi21d = vi21d[nummi]';
helpvector=0||0||0||0;
vi11d=shape(helpvector,2,2);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+3] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma14 */
contribm[nbeta+4] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma2 */
vi21d=0||0;
vi21d = vi21d[nummi]';
helpvector=0||0||0||2*sigma2;
vi11d=shape(helpvector,2,2);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+5] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma23 */
vi21d=0||1;
vi21d = vi21d[nummi]';
helpvector=0||0||0||0;
vi11d=shape(helpvector,2,2);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+6] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma24 */
contribm[nbeta+7] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma3 */
contribm[nbeta+8] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma34 */
contribm[nbeta+9] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma4 */
contribm[nbeta+10] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
end;
else if (di=4) then do;
j=2;
do while (j <= nrow(yobs));
g = exp(psi[1]+yobs[j-1]*psi[2]);
g = g/(g+1);
hij = 1||yobs[j-1];
contribd = contribd-yobs[j]#hij#g#(1-g);
j = j+1;
end;
g = exp(psi[1]+yobs[nrow(yobs)]*psi[2]);
g = g/(g+1);
muc = mui[di];
mup = mui[nummi];
hist = yobs;
vi11 = vi[nummi,nummi];
vi21 = vi[di,nummi];
vi11i = inv(vi11);
ecurr = muc + vi21*(vi11i)*(hist-mup);
contribd = contribd - ecurr#hij#g#(1-g);
/* use derivative wrt beta */
contribm[1:nbeta] = (1-g)#(xi[di,]-vi21*vi11i*xi[nummi,])';
/* use derivative wrt sigma1 */
vi21d=0||0||0;
vi21d = vi21d[nummi]';
helpvector=2*sigma1||0||0||0||0||0||0||0||0;
vi11d=shape(helpvector,3,3);
```

```
vi11d = vi11d[nummi,nummi];
contribm[nbeta+1] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma12 */
vi21d=0||0||0;
vi21d = vi21d[nummi]‘;
helpvector=0||1||0||1||0||0||0||0||0;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+2] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma13 */
vi21d=0||0||0;
vi21d = vi21d[nummi]‘;
helpvector=0||0||1||0||0||0||1||0||0;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+3] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma14 */
vi21d=1||0||0;
vi21d = vi21d[nummi]‘;
helpvector=0||0||0||0||0||0||0||0||0;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+4] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma2 */
vi21d=0||0||0;
vi21d = vi21d[nummi]‘;
helpvector=0||0||0||0||2*sigma2||0||0||0||0;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+5] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma23 */
vi21d=0||0||0;
vi21d = vi21d[nummi]‘;
helpvector=0||0||0||0||0||1||0||1||0;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+6] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma24 */
vi21d=0||1||0;
vi21d = vi21d[nummi]‘;
helpvector=0||0||0||0||0||0||0||0||0;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+7] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma3 */
vi21d=0||0||0;
vi21d = vi21d[nummi]‘;
helpvector=0||0||0||0||0||0||0||0||2*sigma3;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+8] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
/* use derivative wrt sigma34 */
vi21d=0||0||1;
vi21d = vi21d[nummi]‘;
helpvector=0||0||0||0||0||0||0||0||0;
vi11d=shape(helpvector,3,3);
vi11d = vi11d[nummi,nummi];
contribm[nbeta+9] = (1-g)#(vi21d-vi21*vi11i*vi11d)*vi11i*(hist-mup);
/* use derivative wrt sigma4 */
contribm[nbeta+10] = (1-g)#(0-vi21*vi11i*0)*vi11i*(hist-mup);
end;
end;
if (ind=1) then do;
inflvec = contribm //(contribd‘);
end;
else do;
inflvecind = contribm//(contribd‘);
inflvec = inflvec||inflvecind;
end;
ind = ind+1;
end;
delta = inflvec;
return(delta);
finish delta;
```

```
opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,19,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);

/* calculation hessian */

call nlpfdd(maxlik,grad,hessian,"loglik",est);
create result var {est stde};
append;
delta=delta(est);
nbeta=7;
nsigma=10;
npsi=2;
hessianm = hessian[1:(nbeta+nsigma),1:(nbeta+nsigma)];
hessiand = hessian[(nbeta+nsigma+1):(nbeta+nsigma+npsi),(nbeta+nsigma+1):(nbeta+nsigma+npsi)];

/* calculation the C-matrix */

tweedea = j(nbeta+nsigma,nbeta+nsigma,0);
tweedeb = j(nbeta+nsigma,nbeta+nsigma,0);
tweedea[(nbeta+1):(nbeta+nsigma),(nbeta+1):(nbeta+nsigma)]= inv(hessian[(nbeta+1):(nbeta+nsigma),(nbeta+1):(nbeta+nsigma)]);
tweedeb[1:nbeta,1:nbeta] = inv(hessian[1:nbeta,1:nbeta]);
c = 2*delta`*inv(hessian)*delta;
c1 = 2*delta[1:(nbeta+nsigma),]`*(inv(hessian[1:(nbeta+nsigma),1:(nbeta+nsigma)])-tweedea)
* delta[1:(nbeta+nsigma),];
c2 = 2*delta[1:(nbeta+nsigma),]`*(inv(hessian[1:(nbeta+nsigma),1:(nbeta+nsigma)])-tweedeb)
*delta[1:(nbeta+nsigma),];
call eigen(l,v,-c);
cmax = l[1];
print cmax;
hmax = v[,1];
c3 = 2*delta[(nbeta+nsigma+1):(nbeta+nsigma+npsi),]`*inv(hessiand)*
delta[(nbeta+nsigma+1):(nbeta+nsigma+npsi),];
c12 = c-c3;
ci = -vecdiag(c);
c12i = -vecdiag(c12);
c3i = -vecdiag(c3);
c1i = -vecdiag(c1);
c2i = -vecdiag(c2);
subject = 1:nrow(ci);
create c_matrix var {subject ci c12i c3i c1i c2i hmax};
append;
quit;
proc print data=c_matrix;
run;

data MARCmatrix;
set c_matrix;
run;
proc print data=MARCmatrix;
run;
proc means data=MARCmatrix p95 p99 max min;
var ci c1i c2i c12i c3i hmax;
run;

proc print data=MARCmatrix;
where ci>62.6967641;
run;

proc print data=c_matrix;
where c12i>7;
run;

proc print data=c_matrix;
where c12i>2000;
run;

proc print data=c_matrix;
where c2i>2000;
run;

proc print data=c_matrix;
where c1i>2.85;
```

```
run;

proc print data=c_matrix;
where abs(hmax)>0.25;
run;

/*%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
'     FIGURES
/%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

/* Set the graphics environment */
goptions reset=all cback=white width = 5 htitle=10pt htext=12pt;
/* Use the POINTLABEL option to label the plot points */

proc gplot data=MARCmatrix;
plot ci*subject;
symbol1 pointlabel = ("#subject" h=3 font=swiss)interpol=join
value=dot color=black;
title h=1 'Total C_i';
run;
quit;

goptions reset=all cback=white htitle=12pt htext=10pt;
/* Use the POINTLABEL option to label the plot points */
proc gplot data=MARCmatrix;
plot c12i*subject/ haxis=axis1 vaxis=axis2;
symbol1 pointlabel = ("#subject" h=3 font=swiss) interpol=join
value=dot color=blue;
title h=1 'C_i(theta)';
run;
quit;
goptions reset=all cback=white width = 5  htitle=12pt htext=10pt;
/* Use the POINTLABEL option to label the plot points */
proc gplot data=MARCmatrix;
plot c3i*subject/ haxis=axis1 vaxis=axis2;
symbol1 pointlabel = ("#subject" h=3 font=swiss) interpol=join
value=dot color=blue;
title h=2 'C_i(psi)';
run;
quit;

goptions reset=all cback=white border htitle=12pt htext=10pt;
/* Use the POINTLABEL option to label the plot points */
proc gplot data=MARCmatrix;
plot c1i*subject/ haxis=axis1 vaxis=axis2;
symbol1 pointlabel = ("#subject" h=3 font=swiss) interpol=join
 value=dot color=blue;
title h=1 'C_i(beta)';
run;
quit;
goptions reset=all cback=white htitle=12pt htext=10pt;
/* Use the POINTLABEL option to label the plot points */
proc gplot data=MARCmatrix;
plot c2i*subject/ haxis=axis1 vaxis=axis2;
symbol1 pointlabel = ("#subject" h=3 font=swiss) interpol=join
 value=dot color=blue;
title h=1 'C_i(alpha)';
run;
quit;
goptions reset=all cback=white width = 5 border htitle=12pt htext=10pt;
/* Use the POINTLABEL option to label the plot points */
proc gplot data=MARCmatrix;
plot hmax*subject/ haxis=axis1 vaxis=axis2;
symbol1 pointlabel = ("#subject" h=3 font=swiss) interpol=join
 value=dot color=blue;
title h=1 'hmax';
run;
quit;

goptions reset=all cback=white border htitle=12pt htext=10pt;
/* Use the POINTLABEL option to label the plot points */
proc gplot data=MARCmatrix;
plot c2i*c1i/ haxis=axis1 vaxis=axis2;
symbol1 pointlabel = ("#subject" h=2 font=swiss) interpol=join
 value=dot color=blue;
```

```
title h=1 'C_i(alpha) vs C_i(beta)';
run;
quit;

/*%%%%%%%%%%%%%
'
' Global influence analyses
'
'%%%%%%%%%%%%%%%%

%%/
'
' Analysis of subset 1
'
/*%%%%%%%%%%%%%%%%%%%%%*/



/***************************************/
/* (5a) MCAR analysis without subset 1 */
/***************************************/

data data1mnarsubset1;
set IMPI_2016;
if id in (94, 93) then delete;
run;
proc iml;
use data1mnarsubset1;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMCAR;
read all var {est} into mcarest;
initial=mcarest;
create initial var {initial};
append;
nsub=137-2;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;


proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
```

```
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1]=parameters[23];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);

/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
```

```
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:
```

```
/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,23,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
proc print data=result;
run;
```

```
data MCARsubset1;
set result;
run;
proc print data=MCARsubset1;
run;
data MCARmxlik;
set loglik;
run;
proc print data=MCARmxlik;
run;

/***************************************/
/* (5b) MAR analysis without subset 1  */
/***************************************/
data data1marsubset1;
set IMPI_2016;
if id in (94, 93) then delete;
run;
proc iml;
use data1marsubset1;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMAR;
read all var{est} into MARrest;
initial=MARrest;
create initial var {initial};
append;
nsub=137-2;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
```

```
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:2]=parameters[23:24];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);
/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
```

```
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */
```

```
muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,24,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
proc print data=result;
run;
data MARsubsets1;
set result;
run;
proc print data=MARsubsets1;
run;

data MARmaxlik;
set loglik;
run;
proc print data=MARmaxlik;
run;
```

```
/************************************/
/* (3) DIGGLE-KENWARD MODEL : MNAR   */
/************************************/
data data1nmarsubset1;
set IMPI_2016;
if id in (94, 93) then delete;
run;
proc iml;
use data1nmarsubset1;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aNMAR;
read all var{est} into NMARrest;
*initial=mcarest//0.5;
initial=NMARrest;
create initial var {initial};
append;
nsub=137-2;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
```

```
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:3]=parameters[23:25];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);
/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
```

```
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
```

```
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,25,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
data NMARsubset1;
set result;
run;
proc print data=NMARsubset1;
run;
data NMARmxlik;
set loglik;
run;
proc print data=NMARmxlik;
run;

/*%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

'Analysis of subset
'
/*%%%%%%%%%%%%%%%%%%%%%%%%%%*/

/*************************************/
/* (5a) MCAR analysis without subset 2 */
/*************************************/
```

```
data data1mnarsubset2;
set IMPI_2016;
if id in (33, 78, 133) then delete;
run;
proc iml;
use data1mnarsubset2;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMCAR;
read all var {est} into mcarest;
initial=mcarest;
create initial var {initial};
append;
nsub=137-3;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;



proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
```

```
u55=parameters[22];
psi=j(3,1,0);
psi[1]=parameters[23];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);

/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];
```

```
/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
```

```
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,23,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
proc print data=result;
run;

data MCARsubset2;
set result;
run;
proc print data=MCARsubset2;
run;
data MCARmxlik;
set loglik;
run;
proc print data=MCARmxlik;
run;

/*************************************/
/* (5b) MAR analysis without subset 1  */
/*************************************/
data data1marsubset2;
set IMPI_2016;
if id in (33, 78, 133) then delete;
run;
proc iml;
use data1marsubset2;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
```

```
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMAR;
read all var{est} into MARrest;
initial=MARrest;
create initial var {initial};
append;
nsub=137-3;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:2]=parameters[23:24];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */
```

```
xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);
/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
```

```
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
```

```
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,24,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
proc print data=result;
run;
data MARsubsets2;
set result;
run;
proc print data=MARsubsets2;
run;

data MARmaxlik;
set loglik;
run;
proc print data=MARmaxlik;
run;

/************************************/
/* (3) DIGGLE-KENWARD MODEL : MNAR   */
/************************************/
data data2nmarsubset2;
set IMPI_2016;
if id in (33, 78) then delete;
run;
proc iml;
use data2nmarsubset2;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
```

```
create y var {y};
append;
use impi2aNMAR;
read all var{est} into NMARrest;
*initial=mcarest//0.5;
initial=NMARrest;
create initial var {initial};
append;
nsub=137-2;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:3]=parameters[23:25];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
```

```
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);
/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
```

```
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
```

```
opt[1]=1;
opt[2]=5;
con=j(2,25,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
data NMARsubset2;
set result;
run;
proc print data=NMARsubset2;
run;
data NMARmxlik;
set loglik;
run;
proc print data=NMARmxlik;
run;



/*%%%%%%%%%%%%%%%%%%%%%%%%
'
' Anslysis of subset 3
'
/*%%%%%%%%%%%%%%%%%%%%%%*/


/***************************************/
/* (5a) MCAR analysis without subset 3 */
/***************************************/


data data1mnarsubset3;
set IMPI_2016;
if id in (93, 94, 33, 78, 133) then delete;
run;
proc iml;
use data1mnarsubset3;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
```

```
append;
use impi2aMCAR;
read all var {est} into mcarest;
initial=mcarest;
create initial var {initial};
append;
nsub=137-5;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;


proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1]=parameters[23];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
```

```
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);

/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
```

```
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
```

```
opt[1]=1;
opt[2]=5;
con=j(2,23,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
proc print data=result;
run;

data MCARsubset3;
set result;
run;
proc print data=MCARsubset3;
run;
data MCARmxlik;
set loglik;
run;
proc print data=MCARmxlik;
run;

/***************************************/
/* (5b) MAR analysis without subset 3  */
/***************************************/

data data1marsubset2;
set IMPI_2016;
if id in (93, 94, 33, 78, 133) then delete;
run;
proc iml;
use data1marsubset2;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aMAR;
read all var{est} into MARrest;
initial=MARrest;
create initial var {initial};
append;
nsub=137-5;
```

```
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:2]=parameters[23:24];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
```

```
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
pi = constant('PI');
g=j(nsub,1,0);
/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
```

```
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,24,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
```

```
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
proc print data=result;
run;
data MARsubsets2;
set result;
run;
proc print data=MARsubsets2;
run;

data MARmaxlik;
set loglik;
run;
proc print data=MARmaxlik;
run;

/*************************************/
/* (3) DIGGLE-KENWARD MODEL : MNAR   */
/*************************************/
data data2nmarsubset2;
set IMPI_2016;
if id in (93,94,33, 78,133) then delete;
run;
proc iml;
use data2nmarsubset2;
read all var {id sqcd4 transage art pred time} into data;
id = data[,1];
time = data[,6];
treat = data[,5];
timetreat =time#treat;
art=data[,4];
predart= treat#art;
transage = data[,3];
intercept=j(nrow(data),1,1);
create x var {intercept time treat timetreat art predart transage};
append;
y = data[,2];
create y var {y};
append;
use impi2aNMAR;
read all var{est} into NMARrest;
*initial=mcarest//0.5;
initial=NMARrest;
create initial var {initial};
append;
nsub=137-5;
ntime=5;
create nsub var {nsub};
append;
create ntime var {ntime};
append;
quit;

proc iml;
/*     calculating LOGLIKELIHOOD     */
```

```
use x;
read all into x;
print x;
use y;
read all into y;
use nsub;
read all into nsub;
use ntime;
read all into ntime;
use initial;
read all into initial;
g=j(nsub,1,0);

start integr(yd) global(psi,ecurr,vcurr,lastobs);
pi = constant('PI');
g=exp(psi[1]+psi[2]*lastobs+psi[3]*yd);
g=g/(1+g);
integr=g*(1/(sqrt((2*pi)*vcurr)))*exp(-(1/2)*(((yd-ecurr)**2)/vcurr));
return(integr);
finish integr;

nrun = 0;

start loglik(parameters) global(lastobs,vcurr,ecurr,x,z,y,nsub,ntime,nrun,psi);
beta=parameters[1:7];
u11=parameters[8];
u12=parameters[9];
u13=parameters[10];
u14=parameters[11];
u15=parameters[12];
u22=parameters[13];
u23=parameters[14];
u24=parameters[15];
u25=parameters[16];
u33=parameters[17];
u34=parameters[18];
u35=parameters[19];
u44=parameters[20];
u45=parameters[21];
u55=parameters[22];
psi=j(3,1,0);
psi[1:3]=parameters[23:25];
ll=0;
ind=1;
do while (ind<=nsub);

/* select information on one particular patient */

xi = x[(ind-1)*ntime+1:ind*ntime,];
yi = y[(ind-1)*ntime+1:ind*ntime];
mui = xi*beta;
vi = j(ntime,ntime,1);
ui = j(ntime,ntime,1);
ui[1,1] = u11;
ui[1,2] = u12;
ui[1,3] = u13;
ui[1,4] = u14;
ui[1,5] = u15;
ui[2,1] = 0;
ui[2,2] = u22;
ui[2,3] = u23;
ui[2,4] = u24;
ui[2,5] = u25;
ui[3,1] = 0;
ui[3,2] = 0;
ui[3,3] = u33;
ui[3,4] = u34;
ui[3,5] = u35;
ui[4,1] = 0;
ui[4,2] = 0;
ui[4,3] = 0;
ui[4,4] = u44;
ui[4,5] = u45;
ui[5,5] = u55;
vi=t(ui)*ui;
```

```
pi = constant('PI');
g=j(nsub,1,0);
/* yobs & dropout */

if (yi[nrow(yi)]=.) then
do;
dropout=1;
j=nrow(yi);
do while (yi[j-1]=.);
j=j-1;
end;
di=j;
yobs=yi[1:di-1];
end;
else
do;
dropout=0;
yobs=yi;
end;

/* remove intermittent missingness   */

sequence = t(1:nrow(yobs));
miss=yobs||sequence;
i=1;
do while (i<=nrow(miss));
if (miss[i,1]=.) then do;
nr=nrow(miss);
row=(i-1)*2+1//(i-1)*2+2;
miss=remove(miss,row);
miss=shape(miss,nr-1,2);
end;
else
i=i+1;
end;
yobs = miss[,1];
nummi = miss[,2];

/* Calculation of loglik   */

/* er wordt onderscheid gemaakt tussen wel of geen dropout */
if dropout=0 then   /* COMPLETER */
do;
ll1=0;
ll2=0;
j=2;
do while (j <= nrow(yobs));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g = g/(g+1);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;

/* GET ll1 */
/* yobs=yi[nummi]; */
vvi=vi[nummi,nummi];
vvii=inv(vvi);
detvvi=det(vvi);
muiobs=mui[nummi];
ll1=-(ntime/2)*log(2*pi)-0.5*log(det(vvi))-0.5*t(yobs-muiobs)*vvii*(yobs-muiobs);

/* OBTAIN ll */
ll=ll+ll1+ll2;
end;

else do;          /* WITH DROPOUT */
ll1=0;
ll2=0;
ll3=0;

/* calculate ll2 */
if (di=2) then do;
```

```
goto skip;
end;
else  do;
j=2;
do while (j<=(nrow(yobs)));
g=exp(psi[1]+yobs[j-1]*psi[2]+yobs[j]*psi[3]);
g=g/(1+g);
if (g=1) then do;
g=g-0.0001;
end;
ll2=ll2+log(1-g);
j=j+1;
end;
end;

skip:

/* obtain ll1 */

hist=yobs;
mup=mui[nummi];
vi11=vi[nummi,nummi];
vi11i=inv(vi11);
ll1=-((di-1)/2)*log(2*pi)-0.5*log(det(vi11))-0.5*t(hist-mup)*vi11i*(hist-mup);

/* obtain ll3 */

muc=mui[di];
vi11=vi[nummi,nummi];
vi22=vi[di,di];
vi12=vi[nummi,di];
vi21=vi[di,nummi];
ecurr=muc+vi21*inv(vi11)*(hist-mup);
vcurr=vi22-vi21*inv(vi11)*vi12;
lastobs = yi[nummi[nrow(nummi)]];
a = ecurr-2*sqrt(vcurr)||ecurr+2*sqrt(vcurr);

call quad(integraal,"integr",a);
ll3=integraal;
if (ll3=0) then do;
ll3=ll3+0.00001;
end;
ll3=log(ll3);

/* berekening ll */
ll=ll+ll1+ll2+ll3;
end;
ind=ind+1;
end;
loglik=ll;
nrun=nrun+1;
file log;
put nrun;
return(loglik);
finish loglik;

opt=j(1,11,0);
opt[1]=1;
opt[2]=5;
con=j(2,25,.);
call nlpnrr(rc,est,"loglik",initial,opt,con);
*call nlpnms(rc,est,"loglik",initial,opt,con);
print rc;
call nlpfdd(maxlik,grad,hessian,"loglik",est);
print maxlik;
print grad;
print hessian;
inf = - hessian;
covar = inv(inf);
print covar;
var = vecdiag(covar);
print var;
stde = sqrt(var);
print stde;
print est;
```

```
esttrns=t(est);
print esttrns;
zval=esttrns/stde;
pval=2*(1-cdf('normal',abs(zval),0,1));
create result var {est var stde pval};
append;
create hessianmatrix from hessian;
append from hessian;
create covmatrix from covar;
append from covar;
create loglik from maxlik;
append from maxlik;
quit;
data NMARsubset2;
set result;
run;
proc print data=NMARsubset2;
run;
data NMARmxlik;
set loglik;
run;
proc print data=NMARmxlik;
run;
```

Appendix B

Profile plots of the complete data obtained under the de facto hypotheses and simulation results

B.1 Profile plots of the complete data obtained under the de facto hypotheses

This section displays the complete data mean profile plots using the copy difference in reference (in Figure B.1), copy reference (in Figure B.2), and last mean carried forward (in Figure B.3), de facto estimand assumptions about the missing post-deviation data.

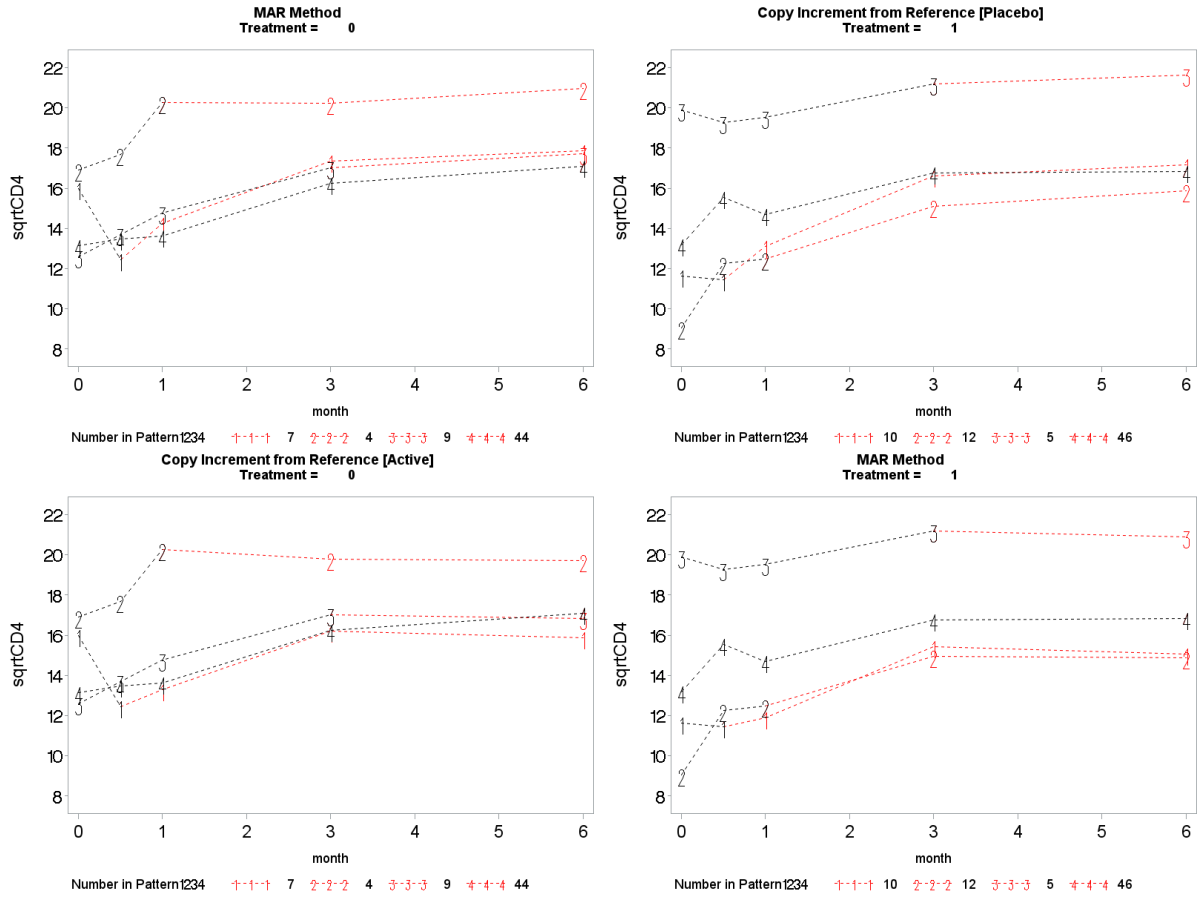

Figure B.1: Copy increment from reference (CDR): placebo arm (top left panel) used as reference to impute data for the active arm (top right panel). Active arm (bottom right panel) used as reference to impute data for the placebo arm (bottom left panel)

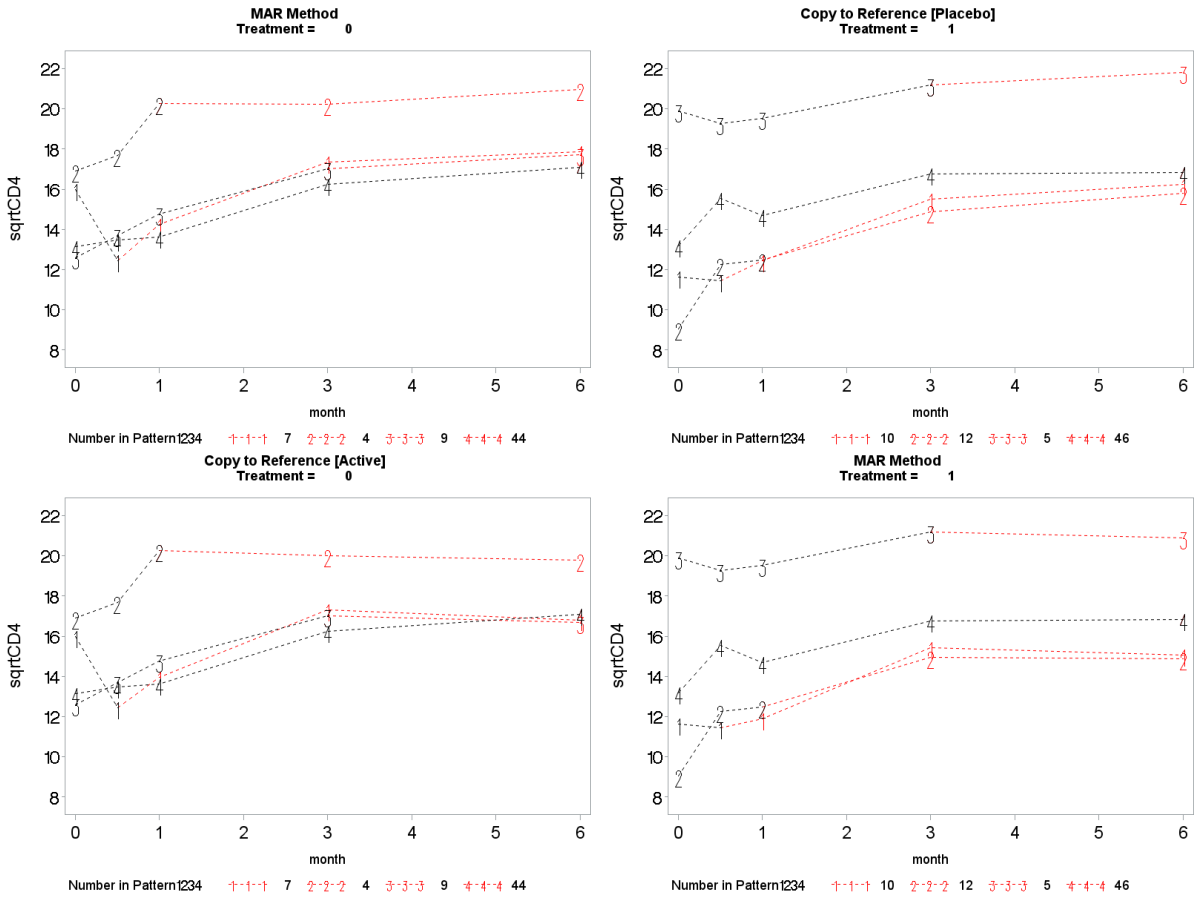

Figure B.2: Copy to reference (CR) : placebo arm (top left panel) used as reference to impute data for the active arm (top right panel). Active arm (bottom right panel) used as reference to impute data for the placebo arm (bottom left panel)

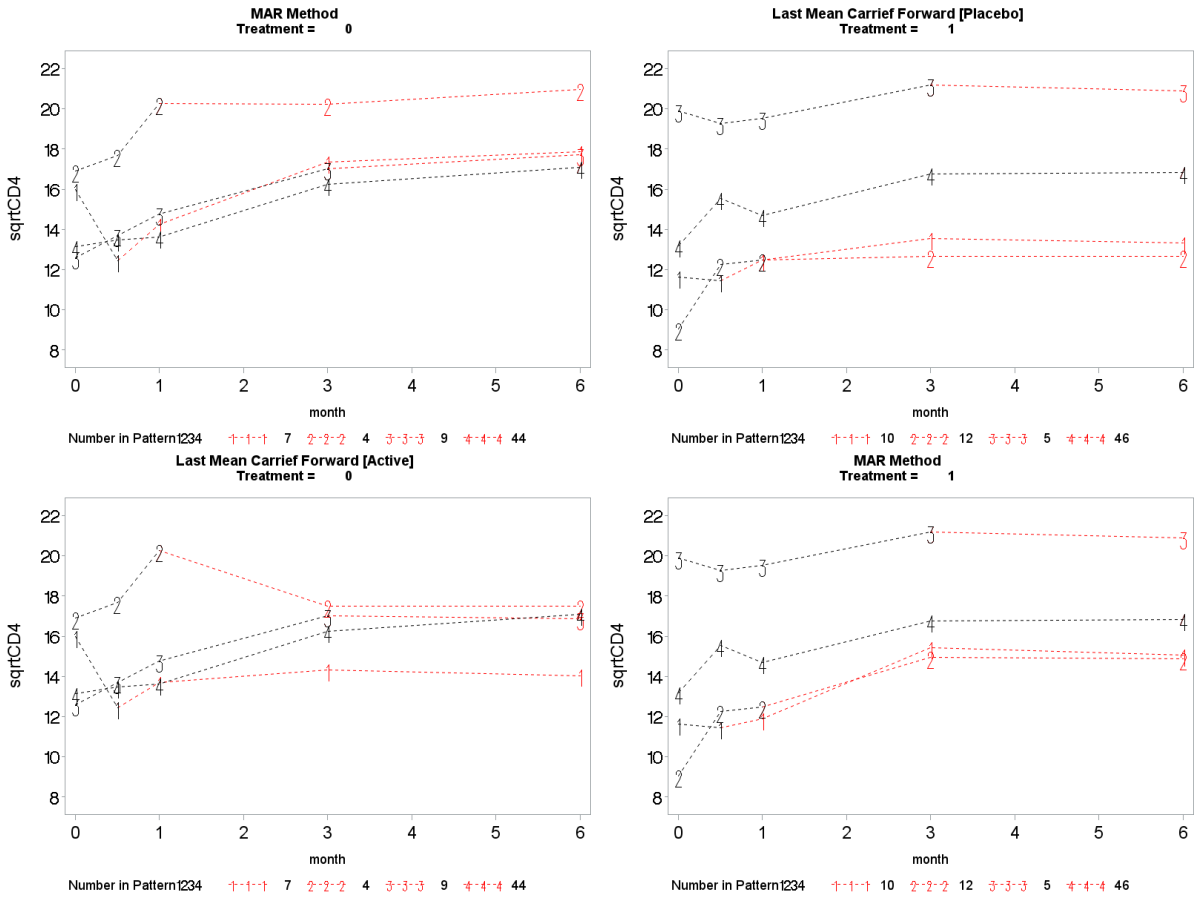

Figure B.3: Last mean carried forward (LMCF): placebo arm (top left panel) used as reference to impute data for the active arm (top right panel). Active arm (bottom right panel) used as reference to impute data for the placebo arm (bottom left panel)

B.2 Simulation results

This section presents simulation results under MCAR and MAR mechanisms with varying missingness rates 5%, 10%, 20%, 30%, and 50% in Tables B.1 and B.2 respectively.

Table B.1: MCAR mechanism by missingness rate

MR[1]		Treatment		Time		Treatment x Time		Coverage probability		
5%	Analysis	Bias	RMSE	Bias	RMSE	Bias	RMSE	Treatment	Time	Treatment x Time
	ML	-0.063	0.127	-0.005	0.015	0.028	0.033	94.10%	96.40%	93.50%
	MI	-0.0602	0.127	0.001	0.020	0.021	0.031	95.50%	94.10%	95.80%
	LMCF	-0.079	0.137	-0.049	0.051	0.032	0.039	93.50%	95.00%	95.60%
	J2R (P-)	-0.100	0.149	-0.006	0.016	0.050	0.054	92.70%	95.40%	94.30%
	CDR (P-)	-0.090	0.141	-0.006	0.016	0.040	0.046	93.10%	95.30%	94.20%
	CR (P-)	-0.095	0.146	-0.006	0.016	0.045	0.050	92.20%	95.30%	94.80%
10%										
	ML	0.112	0.157	-0.017	0,022	0.028	0.035	89.50%	94.50%	95.00%
	MI	0.112	0.158	-0.018	0.024	0.030	0.040	90.50%	95.50%	95.00%
	LMCF	0.095	0.150	-0.103	0.104	0.040	0.046	92.50%	87.50 %	95.50%
	J2R (P-)	0.072	0.133	-0.015	0.020	0.054	0.058	94.5%	87.50%	94.50%
	CDR (P-)	0.088	0.143	-0.015	0.020	0.046	0.051	92.50%	86.50 %	96.00%
	CR (P-)	0.078	0.136	-0.015	0.020	0.052	0.057	93.50%	88.50%	95.50%
20%										
	ML	-0.032	0.12	-0.0024	0.017	0.024	0.034	95.00%	94.50%	95.00%
	MI	-0.042	0.124	-0.005	0.022	0.033	0.045	94.50%	95.50 %	95.00%
	LMCF	-0.099	0.154	-0.183	0.183	0.078	0.083	93.50%	90.50% %	94.50%
	J2R (P-)	-0.122	0.164	0.0002	0.016	0.092	0.100	89.50%	95.00 %	94.50%
	CDR (P-)	-0.074	0.134	0.0002	0.016	0.057	0.063	95.50%	95.00 %	93.50%
	CR (P-)	-0.099	0.149	0.0002	0.016	0.073	0.075	94.50%	95.00 %	94.50%
30%										
	ML	-0.092	0.149	-0.052	0.055	0.043	0.051	94.30%	94.50%	95.10%
	MI	-0.050	136	-0.030	0.039	0.005	0.037	94.40%	95.20%	95.80%
	LMCF	-0.136	0.184	-0.279	0.280	0.081	0.086	89.10%	85.70%	90.20%
	J2R (P-)	-0.220	0.252	-0.059	0.062	0.138	0.141	85.20 %	94.50%	87.40%
	CDR (P-)	-0.160	0.200	-0.059	0.062	0.094	0.098	90.50%	94.10%	95.30%
	CR (P-)	-0.197	0.216	-0.059	0.062	0.109	0.112	90.90%	93.20%	92.70%
50%										
	ML	0.036	0.123	0.0004	0.022	0.009	0.033	96.60%	95.30%	95.40%
	MI	0.039	0.136	-0.006	0.031	0.026	0.051	95.30%	95.10%	94.60%
	LMCF	-0.045	0.129	-0.360	0.360	0.084	0.088	94.30%	89.10%	95.40%
	J2R (P-)	-0.112	0.159	0.0004	0.023	0.127	0.130	89.50%	95.10%	90.60%
	CDR (P-)	-0.045	0.121	0.0004	0.023	0.088	0.089	94.70%	95.20%	93.40%
	CR (P-)	-0.087	0.142	0.0004	0.023	0.112	0. 115	91.20%	95.20%	89.50%

Table B.2: MAR mechanism by missing rate

MR[2]		Treatment		Time		Treatment x Time		Coverage probability		
5%	Analysis	Bias	RMSE	Bias	RMSE	Bias	RMSE	Treatment	Time	Treatment x Time
	ML	-0.153	0.189	-0.0112	0.018	0.009	0.022	87.50%	95.30%	95.10%
	MI	-0.155	0.191	-0.008	0.017	0.004	0.022	86.90%	95.50%	94.50%
	LMCF	-0.155	0.190	-0.033	0.036	0.011	0.023	85.90%	95.10%	95.30%
	J2R (P-)	-0.172	0.205	-0.012	0.019	0.023	0.031	82.50%	94.90%	95.20%
	CDR (P-)	-0.160	0.195	-0.012	0.019	0.015	0.025	87.30%	95.10%	94.50%
	CR (P-)	-0.165	0.199	-0.012	0.018	0.019	0. 027	83.70%	94.30%	95.50%
10%										
	ML	0.062	0.125	0.003	0.015	0.011	0.024	94.50%	95.10%	94.20%
	MI	0.059	0.125	-0.007	0.017	0.016	0.030	94.20%	95.10%	94.30%
	LMCF	0.031	0.11	-0.087	0.088	0.034	0.040	95.10%	93.30%	95.20%
	J2R (P-)	0.013	0.110	-0.004	0.016	0.044	0.050	94.50%	95.20%	96.60%
	CDR (P-)	0.036	0.115	-0.004	0.016	0.032	0.040	95.69%	95.90%	95.20%
	CR (P-)	0.022	0.111	-0.004	0.016	0.040	0. 045	95.30%	95.10%	95.60%
20%										
ML	-0.043	0.012	-0.0061	0.018	0.009	0.025	94.10%	94.50%	95.30%	
	MI	-0.036	0.125	-0.013	0.026	0.013	0.033	95.00%	94.50%	95.30%
	LMCF	-0.127	0.175	-0.180	0.180	0.066	0.070	87.90%	84.40%	94.50%
	J2R (P-)	-0.131	0.177	-0.002	0.017	0.078	0.083	86.80%	95.30%	95.10%
	CDR (P-)	-0.102	0.155	-0.002	0.017	0.054	0.059	85.30%	96.10%	95.50%
	CR (P-)	-0.115	0.163	-0.002	0.017	0.063	0. 067	89.60%	95.10%	95.40%
30%										
	ML	0.062	0.125	0.003	0.015	0.011	0.024	94.50%	95.40%	96.30%
	MI	0.059	0.125	-0.007	0.017	0.016	0.030	95.10%	94.50%	96.20%
	LMCF	0.031	0.11	-0.087	0.088	0.034	0.040	94.30%	95.50%	94.70%
	J2R (P-)	0.013	0.110	-0.004	0.016	0.044	0.050	94.50%	96.10%	95.90%
	CDR (P-)	0.036	0.115	-0.004	0.016	0.032	0.040	95.10%	94.30%	95.50%
	CR (P-)	0.022	0.111	-0.004	0.016	0.040	0.045	94.50%	95.60%	96.20%
50%										
	ML	0.153	0.193	0.055	0.059	-0.071	0.077	87.80%	94.50%	96.10%
	MI	0.144	0.195	0.050	0.059	-0.063	0.077	86.30%	95.40%	96.10%
	LMCF	0.066	0.143	0.374	0.374	0.090	0.094	93.10%	67.80%	92.50%
	J2R (P-)	0.142	0.185	0.054	0.060	0.122	0.125	87.70%	96.10%	89.50%
	CDR (P-)	-0.039	0.124	0.054	0.060	0.056	0.061	95.50%	94.80%	96.40%
	CR (P-)	-0.102	0.156	0.54	0.060	0.095	0.098	93.10%	92.70%	95.40%

Appendix C

SAS code for implementing sensitivity analyses for the shared-parameter model framework using SAS version 9.5

This section presents a SAS code for implementing the sensitivity analyses for to CD4 count measurement data (Impicd4data.sas) from the IMPI trial. The dataset should be in a long format as

```
Data Impicd4data;
input idnum art treat transage time respons resptype;
cards;
1  0 1 33  0     15.26434    1
1  0 1 33  0.5   15.52417    1
1  0 1 33  1     11.7047     1
1  1 1 33  3     16.8226     1
1  1 1 33  6     18.02776    1
1  0 1 33  0     0           0
1  0 1 33  0.5   0           0
1  0 1 33  1     0           0
1  1 1 33  3     0           0
1  1 1 33  6     0           0
.  .    .   .   .    .      .
RUN;
```

Data variables description

idnum: subject identification number.

respons: square root of CD4 count.

transage: age of a subject.

art: anti-retroviral therapy.

treat: prednisolone.

time: month.

The variables "idnum" an indicator for subject, "treat" a (binary) indicator for treatment ("0" for placebo arm, "1" for the prednisolone arm), and "time" time in months (0, 0.05, 1, 3, 6 months). The actual outcome, $\sqrt{\text{CD4}}$ count, and the dropout indicators are stacked into a single variable 'respons', with 'resptype' indicating whether the outcome listed is missing ("1") or observed ("0"). It is sometimes convenient to dispose of character variables indicating the outcome distribution, "dist", and the link function chosen "link", respectively.

```
data Impicd4data;
set Impicd4data;
dist="GAUS";
if resptype=0 then dist="BINA";
link="IDEN";
if resptype=0 then link="LOGI";
run;
data hulp;
set Impicd4data;
if (resptype=0 and time=0) then respons=.;
run;
```

The NLMIXED procedure can then be invoked for parameter estimation and for prediction of unobserved measurements:

```
/* Sen. 1*/
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gammag=1 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0
gamma6=0 sigma=2.5 taug=2;
if ytype=1 then do;
mean = beta0 + g + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + g*gammag + gamma1*treat + gamma2*time + gamma3*treat*time +
gamma4*art + gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random g~normal(0,taug*taug) subject=id;
estimate "taug^2" taug*taug;
estimate "gammag^2*taug^2" gammag*gammag*taug*taug;
estimate "sigma^2" sigma*sigma;
predict (beta0 + g + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;

data toenailc2mnar;
set toenailc2mnar;
mnar=y;
if ytype=0 then mnar=.;
if (y=. and ytype=1) then mnar=pred;
run;
data toenailc2mar;
set toenailc2mar;
mar=y;
if ytype=0 then mar=.;
if (y=. and ytype=1) then mar=pred;
run;
data toenailc2m;
merge toenailc2mnar toenailc2mar;
if ytype=0 then delete;
keep id treat time y mnar mar;
run;

/*Sen. 2 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gammag=1 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0
gamma6=0 sigma=2.5 taug=2 tauh=2.2;
if ytype=1 then do;
mean = beta0 + g + h + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
```

```
end;
else if ytype=0 then do;
eta = gamma0 + g*gammag + gamma1*treat + gamma2*time + gamma3*treat*time +
gamma4*art + gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random g h ~ normal([0,0],[taug*taug,0,tauh*tauh]) subject=id;
estimate "taug^2" taug*taug;
estimate "tauh^2" tauh*tauh;
estimate "gammag^2*taug^2" gammag*gammag*taug*taug;
estimate "sigma^2" sigma*sigma;
predict (beta0 + g + h + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;
data toenailc2mnar;
set toenailc2mnar;
mnar=y;
if ytype=0 then mnar=.;
if (y=. and ytype=1) then mnar=pred;
run;
data toenailc2mar;
set toenailc2mar;
mar=y;
if ytype=0 then mar=.;
if (y=. and ytype=1) then mar=pred;
run;
data toenailc2m;
merge toenailc2mnar toenailc2mar;
if ytype=0 then delete;
keep id treat time y mnar mar;
run;



/*Sen. 3 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gammag=1 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0
gamma6=0 sigma=2.5 taug=2 tauk=2.2;
if ytype=1 then do;
mean = beta0 + g + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + g*gammag + k*1 + gamma1*treat + gamma2*time + gamma3*treat*time +
gamma4*art + gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random g k ~ normal([0,0],[taug*taug,0,tauk*tauk]) subject=id;
estimate "taug^2" taug*taug;
estimate "tauk^2" tauk*tauk;
estimate "gammag^2*taug^2" gammag*gammag*taug*taug;
estimate "gammak^2*tauk^2" 1*1*tauk*tauk;
estimate "sigma^2" sigma*sigma;
predict (beta0 + g + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;
data toenailc2mnar;
set toenailc2mnar;
mnar=y;
if ytype=0 then mnar=.;
if (y=. and ytype=1) then mnar=pred;
run;
data toenailc2mar;
set toenailc2mar;
mar=y;
```

```
if ytype=0 then mar=.;
if (y=. and ytype=1) then mar=pred;
run;
data toenailc2m;
merge toenailc2mnar toenailc2mar;
if ytype=0 then delete;
keep id treat time y mnar mar;
run;
/*Sen. 4 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gammag=1 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0
gamma6=0 sigma=2.5 taug=2 tauj=2.2;
if ytype=1 then do;
mean = beta0 + g + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + g*gammag + j*1 + gamma1*treat + gamma2*time + gamma3*treat*time
gamma4*art + gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random g j ~ normal([0,0],[taug*taug,0,tauj*tauj]) subject=id;
estimate "taug^2" taug*taug;
estimate "tauj^2" tauj*tauj;
estimate "gammag^2*taug^2" gammag*gammag*taug*taug;
estimate "sigma^2" sigma*sigma;
predict (beta0 + g + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;


/*Sen. 5 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0 gamma6=0
sigma=2.5 tauh=2 tauk=2.2;
if ytype=1 then do;
mean = beta0 + h + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + k*1 + gamma1*treat + gamma2*time + gamma3*treat*time + gamma4*art +
gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random h k ~ normal([0,0],[tauh*tauh,0,tauk*tauk]) subject=id;
estimate "tauh^2" tauh*tauh;
estimate "tauk^2" tauk*tauk;
estimate "sigma^2" sigma*sigma;
predict (beta0 + h + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;


/*Sen. 6 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gammaj =1 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0
gamma6=0 sigma=2.5 tauj=2;
if resptype=1 then do;
mean = beta0 + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
```

```
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(respons-mean)**2/(sigma**2);
ll = dens;
end;
else if resptype=0 then do;
eta = gamma0 + j*gammaj + gamma1*treat + gamma2*time + gamma3*treat*time +
gamma4*art + gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = respons*log(p) + (1-respons)*log(1-p);
end;
model respons ~ general(ll);
random j ~ normal(0,tauj*tauj) subject=idnum;
estimate "tauj^2" tauj*tauj;
estimate "gammaj^2*tauj^2" gammaj*gammaj*tauj*tauj;
estimate "sigma^2" sigma*sigma;
predict (beta0 + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;



/*Sen. 7 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gammaj =1 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0
gamma6=0 sigma=2.5 tauh=2 tauj=2.2;
if ytype=1 then do;
mean = beta0 + h + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + j*gammaj + gamma1*treat + gamma2*time + gamma3*treat*time +
gamma4*art + gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random h j ~ normal([0,0],[tauh* tauh,0,tauj*tauj]) subject=id;
estimate "tauh^2" tauh*tauh;
estimate "tauj^2" tauj*tauj;
estimate "gammaj^2*tauj^2" gammaj*gammaj*tauj*tauj;
estimate "sigma^2" sigma*sigma;
predict (beta0 + h + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;



/*Sen. 8 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gamma1=0 gamma2=0 gamma3=0 gamma4=0 gamma5=0 gamma6=0
sigma=2.5 tauh=2;
if ytype=1 then do;
mean = beta0 + h + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + gamma1*treat + gamma2*time + gamma3*treat*time + gamma4*art +
gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random h ~ normal(0,tauh*tauh) subject=id;
estimate "tauh^2" tauh*tauh;
estimate "sigma^2" sigma*sigma;
predict (beta0 + h + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
```

```
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;



/*Sen. 9 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gamma1=0 gammaj = 1 gamma2=0 gamma3=0 gamma4=0 gamma5=0
gamma6=0 sigma=2.5 tauj=2;
if ytype=1 then do;
mean = beta0 + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + j*gammaj + gamma1*treat + gamma2*time + gamma3*treat*time +
gamma4*art + gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random j ~ normal(0,tauj*tauj) subject=id;
estimate "tauj^2" tauj*tauj;
estimate "gammaj^2*tauj^2" gammaj*gammaj*tauj*tauj;
estimate "sigma^2" sigma*sigma;
predict (beta0 + j + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;

/*Sen. 10 */
proc nlmixed data=hulp qpoints=20 maxiter=500;
parms beta0=23.23 beta1= 0.25 beta2=0.42 beta3=-0.09 beta4 = 2.66 beta5 = -0.29
beta6 = -2.94 gamma0= 0.35 gamma1=0.5 gamma2=0.06 gamma3=0.04 gamma4=0.02
gamma5=0.1 gamma6=0.05 sigma=2.5 tauk=2;
if ytype=1 then do;
mean = beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage;
dens = -0.5*log(3.14159265358) - log(sigma) - 0.5*(y-mean)**2/(sigma**2);
ll = dens;
end;
else if ytype=0 then do;
eta = gamma0 + k*1 + gamma1*treat + gamma2*time + gamma3*treat*time + gamma4*art +
gamma5*treat*art + gamma6*transage;
p = exp(eta)/(1+exp(eta));
ll = y*log(p) + (1-y)*log(1-p);
end;
model y ~ general(ll);
random k ~ normal(0,tauk*tauk) subject=id;
estimate "tauk^2" tauk*tauk;
estimate "sigma^2" sigma*sigma;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mnar;
predict (beta0 + beta1*treat + beta2*time + beta3*treat*time + beta4*art +
beta5*treat*art + beta6*transage) out=toenailc2mar;
run;

data toenailc2mnar;
set toenailc2mnar;
mnar=y;
if ytype=0 then mnar=.;
if (y=. and ytype=1) then mnar=pred;
run;
data toenailc2mar;
set toenailc2mar;
mar=y;
if ytype=0 then mar=.;
if (y=. and ytype=1) then mar=pred;
run;
data toenailc2m;
```

```
merge toenailc2mnar toenailc2mar;
if ytype=0 then delete;
keep id treat time y mnar mar;
run;
```