Vol.2.Issue.3., 2014

Bulletin of Pharmaceutical and Medical Sciences (BOPAMS)
A Peer Reviewed International Journal http://www.bopams.com

REVIEW ARTICLE

INTERNATIONAL STANDARD SERIAL NUMBER
E-ISSN 2341-4103

A REVIEW OF METHODS FOR HANDLING MISSING DATA IN THE FORM OF DROPOUTS IN LONGITUDINAL CLINICAL TRIALS

A. SATTY AND H. MWAMBI

School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private Bag X01 Scottsville 3209, Pietermaritzburg, South Africa

Article Info:
Received: 14/09/2014
Revised on : 27/09/2014
Accepted on: 29/09/2014

ABSTRACT

Much clinical trials data-based research is characterized by the unavoidable problem of dropout as a result of missing or erroneous values. This paper aims to review some of the vari-ous techniques to address the dropout problems in longitudinal clinical trials. The fundamental concepts of the patterns and mechanisms of dropout are discussed. This study presents five general techniques for handling dropout: (1) Deletion methods; (2) Imputation-based methods;(3) Data augmentation methods; (4) Likelihood-based methods; and (5) MNAR-based meth-ods. Under each technique, several methods that are commonly used to deal with dropout are presented, including a review of the existing literature in which we examine the effectiveness of these methods in the analysis of incomplete data. Two application examples are presented to study the potential strengthes or weaknesses of some of the methods under certain dropout mechanisms as well as to assess the sensitivity of the modelling assumptions.

Keywords: Incomplete longitudinal clinical trials, Missing at random (MAR), Imputation, Weighting methods, Sensitivity analysis.

INTRODUCTION

Longitudinal clinical studies repeatedly measure the outcome of interest and covariates over a sequences of time points. Longitudinal studies play a vital role in many disciplines of science including medicine, epidemiology, ecology and public health. However, data arising from such studies often show inevitable incompleteness due to dropouts or lack of follow-up. More generally, a patient's outcome can be missing at one follow-up time and be measured at the next follow-up time. This leads to a large class of dropout patterns. This paper only pays attention to the monotone dropout pattern that results from attrition, in the sense that if a patient drops out from the study prematurely, then on that patient no subsequent repeated measurements of the outcome are obtained. Other types of dropout patterns are possible, such as intermittent dropout, but we focus on monotone dropout pattern as it is common in longitudinal studies. These commonly include studies done by the pharmaceutical industry as contained in protocols for many conditions where data are not collected after a study participant discontinues study treatment. This is highlighted in a recent report on the prevention and treatment of dropout by the National Research Council. A summary of the report was provided

by Little et al. (2012). However, even in these studies, there typically is both unplanned and planned dropout. A predominately monotone pattern for missing outcome data is less common in clinical outcome studies and in publically-funded trials which are more of a pragmatic nature (e.g., trials in which the intention-to-treat estimand is the primary objective).

Given the problems that can arise when there are dropouts in longitudinal clinical trials, the following question is forced upon researchers. What methods can be utilized to handle these po-tential pitfalls? The goal is to use approaches that better avoid the generation of biased results.

The choice of statistical methods for handling dropouts has important implications on the esti-mation of the treatment effects, depending on whether one is considering a more of a pragmatic nature analysis or a more exploratory analysis. In case of a pragmatic analysis (intention-to-treat analysis), the goal of the clinical trial researchers is to produce a pragmatic analysis of the data.

However, for incomplete longitudinal clinical trials, the dropouts complicate this process as most of the methods to be used to dealing with the dropout problem produce an exploratory analysis in nature rather than a pragmatic perspective. The literature presents various techniques that can be used to deal with dropout, and these range from simple classical ad hoc methods to model-based methods. These methods should be fully understood and appropriately characterized in relation to dropouts and should be theoretically proven before they are used practically. Further, each method is based on a specific dropout mechanism, but one needs to realize that at the heart of the dropout problems it is impossible to identify the dropout mechanism (will be discussed later). Thus, it is important to address the mechanisms that govern dropouts. It is noted that (Little, 1995) the term dropout mechanism can be used, instead of missing data, when it relates to subjects dropping out of a clinical trial study prematurely, particularly in the context of longitudinal studies.

In this review article, we review some of the various techniques to address the dropout problem in longitudinal clinical trials. The main objective is to investigate various techniques, and to discuss the most appropriate techniques as well as gain insight into the appropriateness of these techniques for handling incomplete longitudinal clinical trials due to dropouts. This paper is divided into five strategies: deletion, imputation-based, data augmentation, dealing with data as incomplete, and MNAR model-based. In Section 2, a necessary notation in terms of the underlying dropout mechanisms is introduced. In Section 3, an overview of methods for analyzing incomplete longitudinal clinical trials data is given with the focus on the aforementioned strategies. In Section 4, two application examples are presented including a description of the full data sets used in the analysis and the study designs. Full analysis and results of the applications are given. Finally, the paper ends with a conclusion in Section 5.

2.**Dropout mechanisms**

A major issue concerning dropout mechanisms is to explain why data are dropping out. Dropout mechanisms however do not imply knowledge about how the dropouts came to be unavailable. The term dropout is misused by many researchers as, in many trials, data are missing not because a participant chooses to drop out but instead because the protocol is written not to follow partici-pants following treatment discontinuation. Discontinuation might be due to adverse effects, lack of efficiency in the execution of the study, both of these reasons, or other reasons. As demonstrated by Rubin (1976), the mechanisms that lead to missingness can be classified into three basic categories. Data are considered missing completely at random (MCAR) when the mechanism that generates the dropouts is a truly random process unrelated to any measured or unmeasured characteristic of the study participants. A second category is missing at random (MAR) in which the dropout mechanism is random meaning, conditional on the observed measurements characteristics of the study sample, the dropout mechanism is independent of the unobserved measurements. Finally, missing not at random (MNAR), is one in which the dropout process depends on unobserved mea-surements and possibly on the observed measurement characteristics of the study sample. Let Yij be the response measurement of individual i at time j, where i = 1, 2, ...N and j = 1, 2, ...n, which can be observed or missing. Let Rij be an indicator variable, where Rij = 1 if Yij is observed and

Rij = 0 if Yij is missing. We now split Yi as Yi = (Yio, Yim), representing observed and unobserved measurements, respectively. Let Di be a dropout indicator for each individual i,

$$D_i = 1 + \sum_{t=1}^{n} R_{ij}. \quad (1)$$

Model (1) measures the occasion when the dropout occurs. The full data for the ith individual are given by Yi and Ri, with joint pdf that can be factorized as

$$f(y_i, r_i \mid X_i, \theta, \gamma) = f(y_i \mid X_i, \theta) f(r_i \mid y_i, X_i, \gamma), \quad (2)$$

where θ is the measurement process, γ is the dropout process and Xi is the design matrix of covariates for the ith individual. The model for dropout process can be re-written as

$$\iota_{id_i} = f(r_i \mid y_i, X_i, \gamma) = Pr(D_i = d_i \mid y_i, X_i, \gamma). \quad (3)$$

where di is a realization of the random variable Di. In Equation (1), it is assumed that all subjects are observed on the first occasion so that Di takes possible values between 2 and n + 1. The maximum value (n + 1) corresponds to a complete measurement sequence. Using Equation (3), the MCAR dropout model reduces to P (Di = di | Yi, Xi, γ) = P (Di = di | Xi, γ), while the MAR dropout model is given by: P (Di = di | Yi, Xi, γ) = P (Di = di | Yio, Xi, γ), where dependence on Yi is only through Yio. Rubin's (1976) classifications are related to the level of bias that dropout may exert on statistical analysis where it is stated that MCAR has negligible potential impact and MNAR has the greatest potential impact. Furthermore, it is impossible to distinguish which underlying mechanism of dropout is in play, unless one knows the motivation for a patients's dropping out. This problem is discussed further in Molenberghs et al. (2008) who show that the formal-based distinction between MAR and MNAR is not possible. This is because for any MNAR model there exists an MAR model that fits the data equally well, but they differ in the prediction of what is unobserved. Hence, it is broadly agreed that the role of such MNAR model is in sensitivity analysis. On the other hand, there are two important broad classes of dropout: dropout that is ignorable from the analysis, and dropout that is non-ignorable. If one can reasonably assume that dropout occur under either MCAR or MAR conditions, the problem is deemed ignorable, and the dropout process need not be explicitly modeled. Moreover, when data are MCAR or MAR, the likelihood-based and Bayesian frameworks allow to ignore the dropout process since they use only observed data, conditional on the model being correctly specified (Little and Rubin, 2002). In contrast, when data are MNAR, the dropout process cannot be ignored from the analysis. In the application to dropout classifications, ignorability, as it applies to dropout mechanisms, does not mean that investigators can ignore dropouts. It refers to the fact that factors that cause dropout are unrelated or weakly related to the estimated intervention effect. In a restricted sense, the term refers to whether dropout mechanisms must be modeled as part of the parameter estimation process or not (Allison, 2002).

3.Techniques for handling dropout Different methods have been suggested for dealing with dropout problems. Different techniques or methods use different approaches to addressing dropout problems. Although the dropout problem is ubiquitous, there is still no firm consensus on what statistical procedures should be used for analysis or on the circumstances under which they should be applied. What follows now is a brief description of the several methods that are commonly used to deal with dropout, including a review of the existing literature in which we examine the effectiveness of these methods in the analysis of incomplete longitudinal clinical trials data.

3.1 Deletion methods

There are several ways to deal with dropout. One of them is to discard patients with incomplete sequences, and then analyze only the units with complete data. Methods that use this approach are called deletion methods. These methods do not replace or impute dropouts and do not make other adjustments to account for dropout. They share many properties in terms of dropout mechanisms and the inefficiencies inherent in losing data for statistical power, although not all to the same degree. The main advantage of these techniques is their simplicity, and the ease with which they can be applied using much of the standard statistical softwares. Brown (1983) states that some of the deletion methods are good options, but only when used

under specific circumstances (i.e., when the amount of dropouts is small and when data are MCAR, for example, the complete case discussed below and the available case which uses all available cases and discards data only at the level of the variable, not at observation level). Namely, both methods need that f(Yio, Di | Xi) = f(Yio | Xi)f(Di | Xi), where f(Yio | Xi) represents the ordinary marginal distribution of the complete data. This implies that f(Yio | Di, Xi) = f(Yio | Xi), that is, the dropout is MCAR. However, because such circumstances are rare, McKnight et al. (2007) advise that one should avoid the deletion methods whenever possible. Furthermore, Little and Rubin (2002) do not recommend any of the deletion methods except in specific situations where the amount of dropouts is limited. Next, we briefly discuss the complete case analysis as a deletion method, explaining its use, strengths and weaknesses.

3.2 Complete case method

The simplest deletion approach is the complete case analysis or list-wise deletion analysis in which the analysis uses only those patients with completely recorded observations. In other words, for all variables under consideration, the complete case confines attention to observations that are available. For example, in longitudinal clinical trials, this method uses only those patients with observed responses at each time point. This method has numerous advantages. The first is its simplicity in that the method can be quite effective and may be satisfactorily used with small amounts of dropouts. However, it is important to make sure that, even in such situations, the deleted cases are not unduly influential (Schafer and Graham, 2002). The second advantage to complete case analysis is that, it is easy to carry out. It is used by default routines in most statistical software packages, but it has varying details of implementations. The primary disadvantages of this method are that: (1) it can produce inefficient estimates, in the sense of loss of statistical power specifically when drawing inferences for sub-populations; and (2) when data are not MCAR, then the method can lead to serious biased results. In other words, it is a valid method only when data are MCAR (Little and Rubin, 1987), but even when MCAR holds, it can still be inefficient (Schafer and Graham, 2002). Thus, McKnight et al. (2007) state that one should give careful consideration before the use of this method regardless of its ease of use. Furthermore, it is easy to envisage situations where complete case can be very misleading. Wang-Clow et al. (1995) presents examples where the complete case has led to misleading results.

3.3 Imputation-based methods

In contrast to the above mentioned technique, we now discuss methods that do generate possible values for the dropouts. These alternative methods are called imputation methods, where one fills-in (imputes) the dropouts to obtain a full data set, and the resultant data are then analyzed by standard statistical methods without concern as if the set represented the true and complete data set (Little and Rubin, 1987). This is the key idea behind commonly used procedures for imputation which include, simple and multiple imputation (M). Simple imputation techniques substitute one value for every dropout in the data set (Little and Rubin, 1987, 2002). In contrast to simple imputation techniques, MI fills in more than one value for each dropout to allow for the appropriate evaluation of imputation uncertainty (Little and Rubin, 1987).

3.4 Simple imputation methods

There are simple imputation methods that include: (1) mean imputation, in which dropouts are replaced with the estimated mean of the data set; (2) last observation carried forward (LOCF), in which every dropped out value is replaced by the last observed value from the patient, i.e., it is a method that assumes that the outcomes would not have changed from the last observed value.

The method assumes that if Yij (endpoint measurement) is missing due to dropout from the study then Yij = Yij−1, where j = 2, ..., J − 1, meaning if the endpoint measurement is dropping out of the trial, it would be imputed by the last observed measurement. LOCF has been recognized as a popular technique in dealing with incomplete longitudinal clinical trials data. It does well when the dropout mechanism is assumed to be MCAR. However, because such a circumstance is rare, Molenberghs and Kenward (2007) advise that one should avoid this method whenever possible. This method will be revisited in detail in the application section. LOCF will be revisited in the application section; (3) regression imputation, where the dropouts are imputed using the prediction taken from a multiple regression analysis; (4) Hot Deck imputation, in which the dropouts can be

replaced with the observed data taken from a matched data from the variables that contain observed values; and (5) stochastic regression imputation, in which dropouts are replaced by a value that is predicted using regression imputation plus a residual that is drawn to reflect uncertainty in the predicted value. Simple imputation methods are general and flexible for handling dropout, and can be implemented quickly in several statistical softwares (for example, SAS, R, S+ and others).

However, with respect to accurately reproducing known population results (parameter estimates and standard errors), each of these single imputation methods have been found to be inadequate. The problems linked with these techniques include: (1) the performance of these techniques is poor even when the ignorable dropout assumption (MCAR or MAR) holds, a situation that limits their suitability to quite a restricted set of assumptions (Allison, 2002); (2) they may produce seriously biased results that may or may not be predictable; (3) when using these techniques, the standard errors and standard deviations tend to be underestimated, and therefore there is a greater likelihood of committing type-I error (see, Schafer and Graham, 2002). The reason is that variability of the estimators is also underestimated since imputed data are treated as observed data; and (4) these techniques may present inconsistent point estimates when data are MCAR.

3.5 Multiple imputation methods (MI)

MI has received a significant amount of attention in the literature recently. This method is a simulation-based approach that imputes dropouts multiple times (Little and Rubin, 1987). MI is valid under the MAR dropout assumption (Little and Rubin, 1987). The key idea of this approach is to fill in the dropouts multiple times in order to construct multiple complete data sets. MI involves three distinct steps: First, the dropouts are filled in M times to generate M complete data sets. In the process of filling in dropouts, a joint distribution for the complete data set (including observed and unobserved data) and a prior distribution of parameters are assumed for the data augmentation algorithm to simulate random draws from the dropout distribution. Under the MAR dropout mechanism, M independent random numbers can, given the observed values, be generated from the stationary conditional distribution of the dropouts, using the Bayesian estimation technique. After the imputation step, M complete data sets are obtained. The use of the number of imputation (M) needs to be specified. Most often, the choice of M=5 is considered adequate and the efficiency of the parameter estimate based on imputation is given by $(1+\xi/M)-1$, where ξ is the rate of dropout (Rubin, 1987). This formula shows that the relative efficiency of the MI inference is related to the dropout rate (ξ) in combination with the number of imputations (M). Rubin's (1987) simulation indicates that the number of imputations can generally be constrained to be fewer than 10. Many statistical practices tend to support Rubin's heuristics of 3 to 10 imputations. Second, each of the M complete data sets are analyzed using standard procedures, such as linear mixed model, depending on the types of response and assumptions used for the model. Third, the estimates from the M analyses are then combined to produce a single estimate that incorporates the usual sampling variability as well as the variability due to the dropouts.

Further, we assume that the vector of repeated measurements Yi is described by the parameter vector β. In the first imputation step, the objective is to impute the dropouts with draws from meaning that in the process, we generate draws from the distribution of β, thus taking sampling uncertainty of estimating β into account. Alternatively, a Bayesian approach in which uncertainty about β is incorporated by means of using some prior distribution for β. After formulating the posterior distribution of β, the following imputation algorithm is used: A random $\beta*$ is first drawn from the posterior distribution of β, then a random Yim is selected from f(yim | yio, $\beta*$). This posterior distribution is approximated by the normal distribution. The so-imputed dropouts are next augmented to the observed data, yielding complete data, Y = (Y o, Y m), which are then used to obtain β and its variance, V = V ar(β). The steps mentioned above are independently repeated a number of times, say M times, yielding $\beta*$m and V m, for m = 1, ..., M. Finally, the results from the M completed (imputed) data are combined into a single inference. The overall estimated parameter for β and its estimated variance V are

with W and B representing the average within-imputation variance and the between-imputation variance, respectively (Rubin, 1987). There is an important question to be solved when applying MI approach, that is

what variables should be included in the imputation model. The MI inference assumes that the model that is used to analyze the multiply imputed data (the analysis model) is the same as the model used to impute dropouts (the imputation model). However, practically, the two models necessarily need not be the same (Schafer, 1997). The quality of the imputation model will influence the quality of the analysis model results, so it is important to carefully consider the design of the imputation model. Therefore, to obtain high-quality imputations for a particular variable, the imputation model should include variables that are potentially related to the imputed variable and variables that are potentially related to the missingness of the imputed variable (Schafer, 1997). Van Buuren et al. (1999) recommended to include the following covariates in the imputation model:

variables in the analysis model, variables associated with dropouts of the imputed variable, and variables correlated with the imputed variable. However, one can also include auxiliary variables, which may or may not have dropouts. Generally, including variables that do not have dropouts are required in the imputation model.

3.6 Data augmentation methods

Data augmentation methods avoid many of the inherent shortcomings of deletion methods. Such methods derive parameter estimates from the available data as well as from either the probability model or an underlying distribution. In contrast to some of the single imputation methods, data augmentation does not replace dropouts. In estimating parameters, this algorithm takes into ac-count the dropouts, the observed data and the relationships between observed data and several underlying assumptions which is to say that parameter estimates from the observed data are aug-mented by the additional information provided by the proposed probability model or underlying distribution. In the context of incomplete data analysis, Maximum Likelihood (ML), Expectation Maximization (EM), Markov Chain Monte Carlo (MCMC) and weighting methods are considered to be augmentation methods. However, as argued by McKnight et al. (2007), the classification of several of these methods as augmentation methods is not clear-cut, specifically for the MCMC, ML and EM methods. The MCMC method has been referred to as an augmentation method within the context of multiple imputation (Allison, 2002). The ML and EM methods have been described as model-based methods by Little and Rubin (1987), while these procedures have also been referred to as data augmentation by Schafer (1997). We now limit our focus to just a few of these methods as data augmentation methods, namely ML, EM and weighting methods.

3.7 Maximum likelihood (ML)

ML was not designed specifically to deal with dropout in the same way as do, for example, LOCF or multiple imputation. ML is an estimation procedure for estimating parameters under different models such as structural equation models (SEM) and ordinary least squares in regression. We discuss the ML as a method for handling missing data. Examples for applying ML to missing data problems can be found in Little and Rubin (2002). Furthermore, in a variety of situations, ML has proven to be an excellent technique for dealing with dropout. When dropout is ignorable (MAR or MCAR), ML does well, and it produces unbiased estimates (Allison, 2002). Therefore, the ML is fairly easy to describe under this assumption. If the assumption is met, ML estimators for missing data produce estimates that have the following desirable properties: unbiased estimates in large samples, estimates that are asymptotically efficient (small standard errors) and satisfy asymptotic normality which is to say that estimates approximate a normal distribution which can then be used to exploit a normal approximation for statistical inference, such as finding confidence intervals and p-values (McKnight et al., 2007). ML can furthermore be implemented in most statistical software including SAS, SPSS, S-Plus and others.

3.8 Expectation maximization (EM)

The EM algorithm was originally proposed by Dempster et al. (1977). It is the process of calculating and imputing a value for each missing variable based on best prediction models. The

EM algorithm is a very general iterative algorithm for ML estimation in dropout problems. This algorithm requires the less restrictive MAR assumption. The key idea behind EM is to deal with the dropout problem and the complications of estimates related to the ML estimation by attempting to solve smaller complete data problems which lead to parameter estimates for the entire data set (missing and complete data). The EM algorithm handles the dropout using the following steps: (1) fill-in the values for dropout by using estimated values generated by ML; (2) estimate parameters based on data in step 1; (3) re-estimate parameters based on the parameter estimates from step 2; and (4) re-estimate parameters based on the re-estimated data from step 3, and so on, iterating the process until the final step converges on a solution that differs by only a little amount from the previous solution. Each iteration of the EM algorithm consists of two steps, namely the expectation step and the maximization step (Dempster et al., 1977). Each step is completed once within each algorithm cycle which is to say that cycles are repeated until a suitable convergence criterion is satisfied. Assume θ0 is an initial parameter vector and θi are the current measurements. In case of dropouts, the expectation step calculates the objective function that is equal to the expected value of the observed measurement log-likelihood, conditional upon the observed measurement and the current parameters Q(θ | θ i) = R `(θ, Yi)f(Yim | Yio, θi)dYim = E[`(θ | Yi) | Yio, θi], i.e., substituting the expected value of Yim, given Yio and θi. For the maximization step θi+1, the parameter vector that maximizes the log-likelihood of the imputed measurements is, Q(θi+1 | θi) ≥ Q(θ | θ i), for all θ. Further theoretical justification of these steps can be found in Little and Rubin (2002). According to Dempster et al. (1977), the fitted parameters (on convergence) are equal to a local maximum of a likelihood function which is the maximum likelihood estimate in the case of a unique maximum. The EM algorithm has two disadvantages: firstly, it is typically very slow to converge, and secondly, it lacks direct provision of a measure of precision for the maximum likelihood estimates. Several proposals have been made to overcome these drawbacks, and we refer to the techniques as provided by Louis (1982) and Baker (1992).

3.9 Weighting methods

As introduced by Zhao and Lipsitz (1992), weighting methods are based on observed values. In this way, after ignoring all the dropouts from the analysis, the remaining observed values are weighted in accordance with how their distribution approximates the full sample or population. The methods employ the weights in order to correct for either standard errors associated with the parameters or the population variability. To derive suitable weights, the predicted probability of each response is estimated from the data for the variable with dropouts. Assuming that dropout time is discrete, j1, j2, ..., jn, define the set of ordered dropout times, with (Di = jn) for those who complete the study. Therefore, weighting methods only require πi = pr(Di = jn) for the study completers. This πi can be calculated sequentially, i.e., πi1πi2...πij−1, where πij = pr(Di > jn | Di ≥ jn) can be computed from the remaining individuals at jn, conditional upon the history of all observed measurements up to jn. Here we notice that the probability for dropout, πi should be estimated from available measurements, using a series of logistic regressions for the πij 's. Overall, the weighting methods are valid given that the model for πi is correctly specified. Generally speaking, weighting methods are a good option under certain circumstances, for example, when a dropout pattern is monotone or is under univariate analysis. In the context of survey data, Rubin (1987) discusses several methods for applying and estimating weights. Under a suitable joint model for the outcome and covariates, these weighting methods are, in many instances, expected to produce results similar to those of multiple imputation (Schafer and Graham, 2002). In the field of biostatistics, Rubin et al. (1995) developed a weighting regression model that requires an explicit model for the dropout mechanism but relaxes some of the parametric assumptions in the data model. In the case of non-likelihood marginal models, the semi-parametric method of generalized estimating equations (GEE) by Liang and Zeger (1986) has been widely applied for handling dropout. However, GEE requires the stronger MCAR mechanism to hold (Liang and Zeger, 1986). This can be seen by the fact that the GEE score function no longer has zero expectation when a MAR mechanism holds. Two subsequent modifications of the GEE method have been proposed to make it valid under the more general MAR condition: weighted generalized estimating equations (WGEE) and multiple imputation based on generalized estimating equations (MI-GEE). Robins et al.

(1995) devised WGEE extending GEE which requires MAR rather than the much stronger MCAR mechanism, but needs the specification of a dropout model with regard to observed outcomes or covariates, in view of specifying the weights. WGEE involves weighting response measurements by their inverse probability of being observed, estimated from some assumed dropout model (Robins et al., 1995). The idea of WGEE is based on the observed responses after weighting them to account for the probability (propensity) of dropout. Early account of WGEE can be found in Robins et al. (1995). MI-GEE denotes a method based on a combination of MI and GEE model analysis. The primary idea of the combination of MI and GEE comes from Schafer (2003).

He proposed an alternative mode of analysis based on the following steps. (1) Impute the missing outcomes multiple times using a full-parametric model, such as a random effects type model. (2) After drawing the imputations, analyze the so-completed data sets using a conventional marginal model, in this case the GEE method. (3) Finally, perform MI inference on the so-analyzed sets of data. As pointed out by Beunckens et al. (2008), MI-GEE comes down to first using the predictive distribution of the unobserved outcomes, conditional on the observed ones and covariates. In terms of the dropout mechanism, in the MI-GEE method, the imputation model needs to be specified.

This specification can be done by an imputation model that imputes the dropouts with a given set of plausible values (Beunckens et al., 2008). Details of this method can be found in Beunckens et al. (2008). Currently, weighting methods can be carried out in most popular packages, such as STATA, SAS and SUDANA. An alternative weighting method for handling dropout in longitudinal clinical trials is inverse probability weighted (IPW) estimating equations, in which complete cases are weighted by the inverse of their probabilities of being observed in order to adjust for dropouts. This method is valid under MAR assumption (Robins et al., 1995), but requires specification of a dropout model in terms of observed outcomes and/or covariates. IPW has been recognized as an attractive approach because it does not require complete specification of the joint distribution of the longitudinal responses but rather is based only on specification of the first two moments. Several methodological research work in the literature (Robins et al., 1995; Robins and Rotnitzky, 1995) have improved the efficiency of IPW. This improvement leads to the doubly robust estimators. The idea of double robustness was developed by Carpenter et al. (2006).

3.10 Likelihood-based MAR methods

Alternative methods that ignore the dropout mechanism in longitudinal studies are the likelihood-based methods of using available data. Namely, These methods are valid under MAR dropout. This likelihood-based analysis is also termed likelihood-based ignorable analysis, or direct likelihood analysis (DL) (Verbeke and Molenberghs, 2005). These methods use the observed data without the need of neither deletion nor imputation. In other words, no additional data manipulation is necessary when a direct likelihood analysis is envisaged, provided the software tool used for analysis is able to handle measurement sequences of unequal length. The major advantage of these methods is their simplicity, that is, they can be fitted in standard statistical software without involving additional programming, using such tools as SAS software, PROCs MIXED, GLIMMIX and NLMIXED. Despite the flexibility and ease of implementation of likelihood-based methods, there are fundamental issues when selecting a model and assessing its fit to the observed data which do not occur with complete data. These methods are sensible under linear mixed models in combination with the assumption of ignorability. Such models give valid inferences under the restrictive assumption of MAR, where the specification of a dropout model is not necessary, and inference is based on the likelihood function conditional on the observed data alone. In other words, when data are MAR, parameters of the measurement process are not involved in the dropout process which is to say that a likelihood based analysis provides valid inferences, with no need to impute, delete, or weight. This means that the parameters of the measurements and dropout processes are distinct, therefore the estimates based on the maximum likelihood can be drawn by maximizing f(Yio | Xi), where f(Yio | Xi) represents the ordinary marginal distribution of the particular subset of Yi determined by Yio. Hence, there is no need to specify the dropout mechanism for the likelihood-based inference, that is, the contribution of f(Di | Yio, Xi) to the likelihood can be ignored. Note that the likelihood-based methods requires the model for f(Yio | Xi) to be correctly specified as well as they need full

distributional assumptions about Yi. Further, under MNAR dropout, the statistical inferences based on these methods that ignore the dropout mechanism may yield biased results. According to Verbeke and Molenberghs (2000), these mixed-effect models permit the inclusion of patients with dropouts at some time points for both dropout patterns, namely monotone and intermittent. For incomplete longitudinal Gaussian outcome, Likelihood-based mixed effects models were proposed by Laird and Ware (1982). When outcomes are of a non-Gaussian longitudinal clinical trials type, the generalized linear mixed model (GLMM) (Breslow and Clayton, 1993) that is typically estimated through maximum likelihood, can be used. In GLMM, the measurement model and the dropout model are both specified, and the inference is based on maximizing the likelihood function, conditional on the observed data as well as the dropout process. Since DL ideas can be used with a variety of likelihoods, in this study we consider the general linear mixed-effects model (Laird and Ware, 1982) as a key modelling framework which can be combined with the MAR dropout assumption. For Yi, the model can be written as follows

$$Yi = Xi\beta + Zibi + \varepsilon i, \quad (8)$$

where bi ∼ N(0, D), εi ∼ N(0, Σi) and b1, ..., bN , ε1, ..., εN are independent. The meaning of each term in equation (8) are described as follows. Yi is the ni dimensional response vector for subject i, containing the outcomes at ni various measurement occasions, 1 ≤ i ≤ N, N is the number of subjects, Xi and Zi are (ni × p) and (ni × q) dimensional matrices of known covariates, β is the p-dimensional vector containing the fixed effects, bi is the q-dimensional vector containing therandom effects and εi is a ni dimensional vector of residual components, combining measurement error and serial correlation. Finally D is a general (q × q) covariance matrix whose (i, j)th element is dij = dji and Σi is a (ni × ni) covariance matrix which generally depends on i only through its dimension ni, i.e., the set of unknown parameters in Σi will not depend upon i. This means marginally Yi~N(Xiβ, ZiDZ$_i^{'}$ +∑i). Thus if we define Vi = ZiDZ$_i^{'}$i + Σi as the general covariance matrix of Yi, then $f(y_{ij}, \beta, V_i) = (2\Pi)^{\frac{-n}{2}}|V_i|^{\frac{-1}{2}} \exp\{-(y_i - X_i\beta)'V_i^{-1}(y_i - X_i\beta)/2\}$ from which a marginal likelihood can be contributed to estimate β. In the likelihood context, when MAR mechanism and mild regularity conditions hold, parameters _ and are independent, and that likelihood based inference is valid (Little and Rubin, 1987). Then, likelihood of interest is based on the factor $f(Y_i^o \mid \gamma)$.

3.11 MNAR-based methods

All the above methods do not however provide an optimal solution to the problem of MNAR dropout. This kind of dropout poses a major complication, in particular in terms of longitudinal clinical trials setting. There are several applications in the literature which argue that it might be necessary to accommodate dropouts in the modelling process, see, for example, Diggle and Kenward (1994) and Little (1993, 1994). In other words, it is argued that one must model the measurement process jointly with a model for dropout which can itself be considered to be of a scientific interest. Arguably, in terms of MNAR dropout, a wholly satisfactory statistical analysis of the data is not feasible, and therefore more careful consideration is necessary with regard to dealing with the MNAR situation. For MNAR dropout, advanced modelling strategies have been developed by modelling the joint distribution of the dropout indicators pattern and the measurements process (including observed and dropouts). As summarized by Verbeke and Molenberghs (2000) there are at least three factorizations possible to model the joint distribution of the measurements and dropout indicators. First of all, there is outcome-dependence factorization, in which dropout indicators are conditioned on the measurements. Secondly, there is pattern-dependence factorization, in which the distribution of the measurements is a mixture of the distribution for individuals of distinct sub-groups as determined by the dropout patterns. Thirdly, there is parameter-dependence factorization, which is conditional on the group of parameters shared by the two components so that the measurements process and dropout indicators are conditional independent. Correspondingly and based on the above-mentioned factorizations, there are thus three kinds of modelling strategies: selection models (SMs), pattern-mixture models (PMMs) and shared parameter models. According to Verbeke et al. (2001), the practical limitation of any of these model factorizations is that they are sensitive to the assumptions made on the measurements model and the dropout mechanisms. Molenberghs et al. (2004) state that different analysis models can have a distinct impact on

conclusions drawn from the same study. This is the key idea behind commonly used procedures for sensitivity where, given a practical data set, various modelling frameworks with different dropout mechanisms are applied to the same data. Based on local influence methods (Cook, 1986), Verbeke et al. (2001) presented global and local influence methods as additional approaches to study sensitivity. Robins et al. (1998) and Forster and Smith (1998) provided a Bayesian sensitivity analysis routes as alternative frameworks for sensitivity analyses. A more extensive case study on the advantages and disadvantages of several sensitivity routes is not clear-cut as this point is an active area of ongoing research.

3.11.1 Selection models (SMs)

SMs factor the joint distribution of the measurement "complete data", Yi and Di in terms of f(Yi, Di | Xi) = f(Yi | Xi)f(Di | Yi, Xi), i.e, a marginal measurement model that describes the distribution of the underlying complete data, and a dropout mechanism that describes the distribution of the dropout indicators, conditional upon the complete data. In the framework of the SMs, it is not always reasonable to assume that MAR holds, and a wide range modeling approaches for MNAR data have been proposed. One such is the model proposed by Diggle and Kenward (1994) for continuous outcomes with dropout. Diggle and Kenward (1994) consider a SM for the study of a longitudinal measurement when data are MNAR by letting the probability of dropout depend on the possibly unobserved measurements. Similarly, we consider the measurement model to be of the linear mixed effects model (Laird and Ware, 1982), mentioned in equation (8). In agreement with notation introduced in Section 2, the selection model arises when the joint likelihood of the measurement process and the dropout process is factorized as follows

$$f(y_i, r_i \mid X_i, S_i, \theta, \gamma) = f(y_i \mid X_i, \theta) f(r_i \mid y_i, S_i, \gamma), \quad (9)$$

where where Xi denotes the design matrix for fixed effects, Si denotes the design matrix for ran-dom effects. The model for dropout process is based on a logistics regression for the conditional probability of dropout at occasion j, given the subject is still in the study. Assume gi(yij , hij) denote this probability, where hij represent the history of the measurement process. Thus, one can assume that gi(yij , hij) satisfies the model

$$logit[g(h_{ij}, y_{ij})] = logit[p(D_i = j \mid D_i \geq j, h_{ij}, y_{ij})] = \eta(h_{ij}, y_{ij}), \quad (10)$$

where η(hij , yij) is the linear predictor depending on hij and yij . Modelling the dropout mechanism may be simplified in the expression in equation (10) by assuming η(hij , yij) depends only on the current measurement and the previous measurement yij−1, but not on future measurements or higher order history, with corresponding regression coefficients, γ1 and γ2. Higher order history can be modeled, but we assume first order history for simplicity. This leads to the following logistic expression

$$logit[g(y_{i,j-1}, y_{ij})] = logit[p(D_i = j \mid y_{i,j-1}, y_{ij})] = \gamma_0 + \gamma_1 y_{i,j-1} + \gamma_2 y_{ij}. \quad (11)$$

Model (11) contains special cases corresponding to MAR and MCAR mechanisms that can be obtained from γ2=0 or γ1=γ2=0, respectively. A likelihood ratio test (LRT) can be used to compare model fit under a model that assumes the missing data due to dropout are MCAR versus MAR, that is, the LRT for MCAR versus MAR has an approximate χ21 distribution (Diggle and Kenward, 1994). However, the use of the LRT statistic is inappropriate for hypothesis test for MNAR versus MAR when all the other modeling assumptions hold, due to the fact that the behavior of the LRT statistic for the MNAR parameter γ2 is non-standard, since the availability of the information on γ2 is very rare and interwoven with other features of both measurement and dropout models (Jansen et al., 2006). In practice, such a distinction (MAR/MNAR) can only be made using untestable modeling assumptions such the distributional form, see, Kenward (1998). This problem is really laid bare in Molenberghs et al. (2008) which shows that the formal-based distinction between MAR and MNAR is not possible as for any MNAR model there exists an MAR model that fits the data equally well. The similarity of the MAR and MNAR models with respect to fitting to the observed data, may present different predictions of the

unobserved outcomes, conditional upon the observed ones. Hence, for such MNAR models, one recommendation is to use a sensitivity analysis; that is, if the assumptions are changed, the conclusions from the primary (typically MAR) analysis are also changed.

3.11.2 Pattern mixture models (PMMs)

In contrast to SMs, the PMMs specify the joint distribution in terms of f(Yi, Di | Xi) = f(Di | Xi)f(Yi | Di, Xi). PMMs stratify patients according to their dropout pattern. Namely, a separate model is fit for each pattern and then the results can be combined across the different patterns in order to derive an average estimate of the model parameters. Therefore, the joint distribution of the longitudinal measurements as well as the dropout indicators is divided into response pattern so that the distribution of the longitudinal measurements depends on the pattern of responses. The PMMs are under-identified, or possess non-estimable parameters. Therefore, some identifying constraints are required. Little (1993, 1994) proposed the use of the identifying restrictions in which inestimable parameters of the incomplete patterns are set equal to (functions of) the parameters describing the distribution of the completers to deal with under-identifiability of these models. In fact, there is an alternative major strategy simplified to deal with the under-identifiability of PMMs, called model specification in which the different pattern allows for sharing of certain parameters so that the missing pattern can borrow information from patterns with more data points (Verbeke and Molenberghs, 2000). The advantage of this strategy is that the number of parameters decreases which is in general an issue with PMMs. In this article however we focus on the PMMs via an idea of identifying restrictions strategy. Assume that there are t = 1, ..., T dropout patterns, where the dropout indicator, introduced in section 2, is d = t + 1. The complete data density, for pattern t, can be expressed as

ft(y1, ..., yT) = ft(y1, ..., yt)ft(yt+1, ..., yT | y1, ..., yt). (12)

In this equation (12), the first factor ft(y1, ..., yt) is identified from the observed data assuming the first factor is known, and modeled using the observed data. Whereas the second factor is not identifiable from the observed data. In order to identify the second component, the identifying restriction can be applied (Verbeke and Molenberghs, 2000). It is often necessary to base this identification on all patterns for which a given component is identified. We denote this component by ys. Thus, this can be described as

$$f_t(y_s \mid y_1, ..., y_{s-1}) = \sum_{j=s}^{T} \omega_{sj} f_j(y_s \mid y_1, ...y_{s-1}), \qquad s = t+1, ..., T. \tag{13}$$

We denote the set of ωsj used by the vector ωs, components of which are typically non-negative. Every ωs that sums to 1 provides a valid identification scheme. Hence, by incorporating equation (13) into (12), we have

$$f_t(y_1, ..., y_T) = f_t(y_1, ..., y_t) \prod_{s=0}^{T-t-1} \left[\sum_{j=T-s}^{T} \omega_{T-s,j} f_j(y_{T-s} \mid y_1, ..., y_{T-s-1}) \right] \tag{14}$$

To establish the complete data density, it is clear in equation (14) whose information can be used to complement the observed data density in pattern t. According to Little (1993), there are three sets of identifying restrictions associated with such choices of ωs. (1) Complete case missing values (CCMV) when

ft(ys | y1, ..., ys−1) = fT (ys | y1, ..., ys−1), s = t + 1, ..., T, (15)

corresponding to ωsT = 1 and all others equal 0, which is to say that identification is always done from the completers's pattern. (2) Neighboring case missing values (NCMV), in which the nearest identified pattern can be used as follows

ft(ys | y1, ..., ys−1) = fs(ys | y1, ..., ys−1), s = t + 1, ..., T. (16)

The NCMV restriction follows from setting ωs = 1 and all others equal 0. Finally, available case missing values (ACMV). With regard to the corresponding ωs for ACMV, there always is a unique choice. Molenberghs et al. (1998) state that the corresponding ωs can have the following components

$$\omega_{sj} = \frac{\alpha_j f_j(y_1, \ldots, y_{s-1})}{\sum_{\ell=s}^{T} \alpha_\ell f_\ell(y_1, \ldots, y_{s-1})}, \qquad j = s, \ldots, T, \tag{17}$$

where αj is the fraction of observations in pattern j. Clearly, ωsj defined by (17) contains positive components and sum to 1. That is, a valid density function is defined. The SMs and PMMs can be connected using this MAR-ACMV link. The ACMV is reserved for a counterpart of MAR in the PMMs. A specific counterpart to MNAR-SMs has been studied by Kenward et al. (2003).

4 Application examples

4.1 Objectives of the applications

Note that the present study is essentially a simple application scenarios for comparison of meth-ods, rather than an extensive simulation study. This application section has been divided into two application examples: the first example placed strong emphasis on ignorable dropout area which is often used synonymously with MCAR and MAR. In this regard, the scope of the first example was limited to MAR dropout rather than the much stronger assumption MCAR as, in principle, the MCAR assumption is too strong to generally hold in longitudinal clinical trials (Molenberghs and Kenward, 2007). The aim of this example is to investigate some of the aforementioned methods to specify the most appropriate method as well as gain insight into the appropriateness of these tech-niques for handling dropout. The second example focused on MNAR modelling that can be used to deal with the change over time in the outcome score and factors that influence this change in modelling incomplete longitudinal data with continuous outcomes. The aim of the second example is to deal with MNAR dropout by explicitly modelling the assumptions that caused the dropout and incorporated this additional model into the model for the measurement data, and to assess the sensitivity of the modelling assumptions.

4.2 Example 1: The heart rate data

The data set to be analyzed in this example originates from the clinical trial to study the effect of three treatments on heart rate of humans. Full details of this experiment are given in Millikin and Johnson (2009). It is an experiment involving three drugs (AX23, BWW9, and CTRL) and where each patient was measured repeatedly at four different time points (j = 1, 2, 3, 4). After the drug was administered, each patient's heart rate was measured every five minutes for a total of four times.

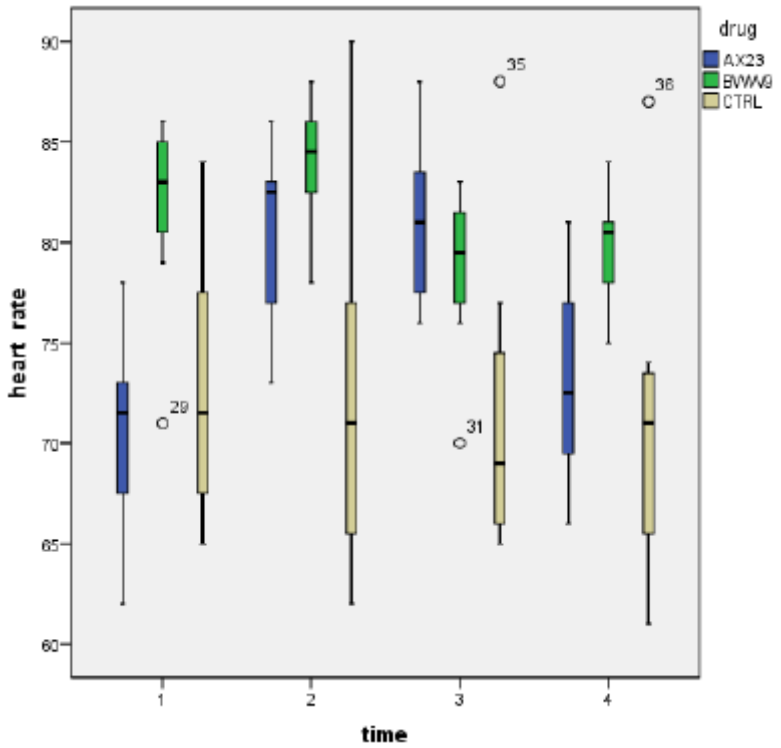

Figure 1: Box-plot for the distribution of heart rate across all four time points for all three drug groups.

To be precise, each patient's heart rate was measured 5, 10, 15 and 20 minutes after administering the treatment. This experiment illustrates the layout for a simple repeated measures experiment. The large size experimental unit is the subject, and the smaller size experiment unit is the time interval when using the split-plot in time notation. At the start of the study n female human subjects were randomly assigned to each drug. Figure 1 shows the distribution of measurements in terms of box-plots at all four time points by all three drug groups. The objective of this experiment was to investigate the drug-response effects, i.e. if the drugs have an effect on heart rate, compare drug groups with each other, including time effects and to find the least-square means. In this paper, we consider the significance of drug main effects, time main effects and the interaction of time and drug effects. The outcome of the analysis reported here was heart rate (HR). Let HRijk denote the heart rate of patient i where i=1,...,8, at time j for j=1,...,4, on drug k, where k=1, 2, 3. We consider the following linear model for HRijk, where the response of the subject i at time j:

$$HRijk = \beta 0 + \beta 1Tj + \beta 2Dk + \beta 3(T * D)jk + \varepsilon ijk, \quad (18)$$

where T is the time, D is the drug, $(T * D)$ is the drug-by-time interaction and εijk are unknown independent and identically distributed normal random error, with mean 0 and variance $\sigma\varepsilon 2$. In this clinical trial, there are no actual dropouts. This provides us with an opportunity to generate MAR dropouts in order to compare the performance of the MI, LOCF, CC and direct likelihood analysis (DL) methods to deal with MAR dropouts. To do so, the following steps were planned and executed:

- We fit a linear mixed model (LMM) (18) using the complete data set to derive parameters of interest.
- From the complete data set, we draw 1000 random samples of n=96. The dropouts were created in the outcome, HR, under 10%, 20% and 30% rates, according to the MAR assump-tion, assuming the dropouts in HR is related to observed values, in the sense that patients with higher heart rate at one measurement occasion tend to drop out of the experiment at the next occasion. Namely, dropouts were created in HR by randomly deleting all observa-tions greater than 75 as a threshold indicating high heart rate. The other

covariates were however kept intact. The observations that triggered the dropouts were kept but all other subsequent observations were deleted. This scenario was generated or replicated 1,000 times. The implication of the MAR assumption in our case is that, patients who are observed to be weaker (deduced by way of their previous observed outcome) are more likely to dropout when they reach a certain value of the HR, as long as their probability of dropout does not further depend upon their missing responses. We assumed a monotone dropout pattern, which is to say that for each patient, if a HR's observation was dropped out for a third time point, the subsequent observation in the fourth time point for that patient was also deleted.

- The MI, DL, CC and LOCF methods were applied to each generated data set. (1) MI was carried out using SAS PROC MI to fill in all the dropouts for each generated sample.

The imputation model is based on model (18) which assumes multivariate normality of the variables. To increase the efficiency of imputation, we used all the available data including the outcome of interest, HR, to predict the dropouts as they are potentially related to the imputed variable as well as to the missingness of the imputed variable. MI was applied to generate M = 5 complete data sets. This is often sufficient to obtain satisfactory results (Rubin, 1987; Schafer, 1999). LMM was then fitted to each imputed data set using SAS procedure MIXED to estimate the overall parameters and their variances. The analysis model that we considered is based on (18). The results of the analysis from these 5 completed (imputed) data sets were combined into a single inference. This was done by using SAS procedure MIANALYZE. (2) LOCF and CC were conducted by using SAS macros available from the authors. After applying LOCF and CC, the same model (18) as before being fitted is analyzed. LOCF replaced the dropouts by the last available observed data, and once the data set has been completed in this way, it is analyzed as if it were fully observed. CC discarded patients with dropouts, and then analyzed only patients with fully observed data.

(3) For DL implemented with PROC MIXED, the data in model (18) was analyzed as they are, consistent with ignorability assumption. Parameters were estimated using Restricted Maximum Likelihood with the Newton-Raphson algorithm.

- Finally, the results from these four methods were then compared with those obtained from the complete data. The comparisons were evaluated in terms of two statistical criteria: bias and efficiency. These criteria are recommended by Schafer and Graham (2002). We defined bias as the difference between the average estimate and the true value. Thus, a better technique is that which does on average approach the population value with less bias. The efficiency criteria has been defined as the variability of the estimates around the true population coef- ficient. Efficiency has been calculated by the average width of the 95% confidence interval.

Thus, a narrower interval is always desirable as it leads to more efficient method.

4.3 Results and discussion from example 1

Table 1: Bias and efficiency of the MI, DL, CC and LOCF methods, under different MAR dropout rate: MIXED least squares means - (interaction terms are not shown)

		Bias				Efficiency			
dropout rate	parameters	MI	DL	CC	LOCF	MI	DL	CC	LOCF
10%	AX23	0.27	0.28	**1.08**	**0.83**	0.83	0.89	**1.09**	**0.93**
	BWW9	-0.18	-0.18	**-1.17**	**-0.25**	0.83	0.90	**1.13**	**0.92**
	CTRL	0.27	0.29	**-1.05**	**0.31**	0.83	0.89	**1.12**	**0.91**
	$time_1$	0.50	0.50	**0.87**	**0.64**	0.97	0.96	**1.09**	**1.00**
	$time_2$	0.50	0.50	**1.55**	**0.78**	0.97	0.96	**1.13**	**0.99**
	$time_3$	-0.01	-0.08	**-1.11**	**-0.15**	0.97	1.09	**1.81**	**1.11**
	$time_4$	0.39	0.49	**1.06**	0.49	0.97	1.07	**1.32**	**1.09**
20%	AX23	0.25	0.41	**2.04**	**0.96**	0.84	0.93	**1.41**	**0.99**
	BWW9	0.50	0.38	**2.05**	**0.97**	0.84	0.94	**1.63**	**1.02**
	CTRL	0.65	0.64	**1.44**	**0.98**	0.84	0.94	**1.46**	**1.02**
	$time_1$	0.08	0.08	**1.19**	**0.87**	0.98	0.96	**1.32**	**1.01**
	$time_2$	0.48	0.48	**2.16**	**0.84**	0.98	0.99	**1.87**	**1.03**
	$time_3$	1.22	1.10	**1.99**	**1.41**	0.98	1.27	**2.14**	**1.56**
	$time_4$	0.06	0.24	**1.41**	**0.93**	0.98	1.27	**1.63**	**1.54**
30%	AX23	0.57	1.24	**2.64**	**2.01**	0.86	1.08	**2.52**	**1.16**
	BWW9	0.89	1.14	**2.10**	**1.94**	0.86	1.08	**2.80**	**1.22**
	CTRL	0.73	1.13	**1.89**	**1.89**	0.86	1.09	**1.93**	**1.34**
	$time_1$	0.56	0.56	**1.41**	**1.39**	1.01	0.97	**1.63**	**1.49**
	$time_2$	0.71	0.71	**2.70**	**1.21**	1.01	0.98	**2.07**	**1.42**
	$time_3$	1.16	1.05	**2.61**	**1.55**	1.01	1.55	**2.68**	**2.11**
	$time_4$	1.20	1.64	**2.88**	**2.13**	1.01	1.58	**2.96**	**1.98**

Note: The largest bias and less efficiency for each given estimate presented in bold. MI=multiple imputation; DL=direct likelihood; CC=complete case; LOCF=last observation carried forward.

The results of MI, DL, CC and LOCF in terms of bias and efficiency, under three dropout rates are listed in Table (1). By looking at this table, we find the following observations. Overall, the MI and DL methods yielded equally good performance and outperformed the CC and LOCF methods. For all dropout rates, among the four methods examined here, CC and LOCF were notable for consistently providing the most biased estimates versus those obtained by MI and DL. The benefits of MI and DL over the CC and LOCF methods are clearly evident. This conclusion was unsurprisingly as both methods, MI and DL, are Bayesian and likelihood based analyses, therefore valid under the MAR dropout assumption (Verbeke and Molenberghs, 2000). The findings based on both MI and DL method were generally similar for all the dropout rates, and in some cases they yielded the same estimates. We refer here to estimates of time1 and time2. Such results should be expected considering the fact that the first and second time points contained observed data for all patients that were considered in the trial. Further, our results confirm the argument put forward by Molenberghs and Kenward (2007), that is; the LOCF method makes the strong assumption that there is no change in the patient response between the observed time points and the missing time period, which can lead to biased estimates. Moreover, the results support the Little and Rubin (2002)'s recommendation to not use CC for handling dropout as it leads to reduction in the sample size which reduces the precision of estimates and therefore can lead to biased results.

As mentioned earlier, a wider interval implies a less efficient, and therefore the widest also implies the worse, 95% is highlighted. An examination of the efficiency condition revealed that, regardless of the dropout rate, the intervals yielded by MI and DL methods were less wide than those based on the CC and LOCF methods, and therefore such estimates were more efficient than were those for CC and LOCF. The CC and LOCF methods were less efficient most frequently, regardless of the dropout rate. This is to be expected for both methods as intuitively the LOCF's weakness is that it tends to create inflated artificial information than truly expected as

imputed values are handled as observed values (Satty and Mwambi, 2012). Further, this CC disadvantage is well documented in Molenberghs and Kenward (2007) who noted that this method can lead to serious inefficient estimates and therefore can be very misleading. On the other hand, between the two CC and LOCF methods, the LOCF was more efficient than the CC method. Efficiency by CC and LOCF methods appeared to be independent of the dropout rate. Overall, the results show that both MI and DL methods offer high efficiency under the MAR dropout mechanism.

4.4 Example 2: The serum cholesterol data

This data concerns the analysis of repeated measures designs and demonstrates how to investigate a specific scenario based on dealing with longitudinal data that has an MNAR dropout mechanism.

The data used here is described and reported in Schoenfield and Lachin (1981). In this trial, 103 patients were randomly assigned to three treatment groups corresponding to two doses; that is, high-dose (750 mg per day), low-dose (375 mg per day) and placebo, and were to be treated for four weeks. This paper is based on a subset of the data on patients who had floating gallstones and who were assigned to the high-dose and placebo groups. In this experiment it was suggested that chenodiol would dissolve gallstones but in doing so might increase levels of serum cholesterol. As a result, serum cholesterol (mg/dL) was measured at baseline and at 6, 12, 20 and 24 weeks of follow-up. Further, many cholesterol measurements contain dropouts because of missed visits, laboratory specimens were lost or inadequate, or patient follow-up was terminated. Additionally, all patients have observed values at time 6. One group of patients received study treatment (drug and placebo), but dropped out of the study before the scheduled post-baseline time. These patients dropped out of the study at time point 12. However, other patients dropped out of the study either at time point 20 or 24. Thereby, the data presents three possible dropout patterns (dropout at time points 12, 20, or 24). All 103 patients are observed at the first occasion, whereas there are 93, 78and 67 patients seen at the second, third and fourth weeks, respectively. The percentage of patients that are still in the study after each week is tabulated in Table 2 by treatment arm. The aim of this clinical trial was to study the safety of the drug chenodiol for the treatment of cholesterol gallstones. However, we restrict our attention to explore the potential influence of dropout on the outcome of interest, the serum cholesterol, as well as the interactive effect of dropout with week and treatment-related influences on the serum cholesterol by using both SMs and PMMs. Findings from the PMMs will be analogous to those from the SMs to obtain additional insights into the serum cholesterol data.

Table 2: NCGS data: Percentage of patients still in study, by treatment arm

week	drug	placebo
6	100	100
12	45	62
20	57	63
24	46	69

Note: Drug=high-dose (750 mg per day)

4.4.1 SMs applied to the serum cholesterol data

We fit the Diggle and Kenward model in accordance with the MCAR, MAR and MNAR assump-tions to the serum cholesterol data. To do so, we combine the measurement model with the logistic regression for dropout model. We assume different intercepts and treatment effects for each of the four time points, with a (4×4) unstructured variance-covariance matrix. In particular, we consider a multivariate normal model, with unconstrained time trend under placebo and an occasion-specific treatment effect. Since the above data contains 103 patients (i=1,...,103) on four time points (j=6, 12, 20 and 24), the fitted model can be expressed as follows

$$Yij = \beta j1 + \beta j2\ i + \varepsilon ij\ , \qquad (19)$$

where i = 0 for placebo and 1 = 1 for active drug. Using this way, we can obtain the parameter estimates and standard errors as well as p-values for the eight mean model parameters. Model (19) was fitted using SAS procedure MIXED with REPEATED statement. For the dropout model, the probability of serum cholesterol is

assumed to follow the logistic regression model. Therefore, we use the following logistic regression model for the dropout model probabilities

logit[g(yij−1, yij)] = logit[p(Di = j | yij−1, yij)] = γ0 + γ1yij−1 + γ2yij , j = 2, 3, 4, 5, (20)

where γ0 is the intercept, γ1 is the effect of the measurement prior to dropout and γ2 is the effect of the measurement at the time of dropout. Model (20) was fitted with an intercept, an effect for previous outcome and an effect for the current unobserved measurement, corresponding to MCAR, MAR and MNAR, respectively. Note that dependence on future unobserved measurements is theoretically possible, but for simplicity, we model dependence on the current unobserved measure-ments. The parameters in Model (20), were estimated using a code written in SAS provided by Dmitrienko et al. (2005) that maximizes the log-likelihood for the model using PROC IML.

4.4.2 PMMs applied to the serum cholesterol data

We now fit the PMMs to the serum cholesterol data using the the idea of identifying restrictions strategy. To do so, we use the following steps: (1) we fit the initial model to the observed data within each of the patterns

ft(y1, ..., yt), (21)

where t = 1, ..., T represent the observed dropout times in the data set. Namely, we fit a sepa-rate model within each pattern. Thereafter the resulting parameter estimates and their estimated variance-covariance matrices were used to extrapolate the patterns. (2) we select an identifica-tion scheme to determine the conditional distributions of the unobserved measurements, given the observed ones

ft(yt+1, ..., yT | y1, ..., yt). (22)

As stated earlier, each of such conditional distributions is a mixture of known normal densities for continuous repeated measures. According to the weights ws introduced in equation (13), an easy way to simulate values from the mixture distribution is to randomly select a component of the mixture and then draw from it. In this regard, we choose an identifying restriction, mentioned earlier, to define the conditional distributions of the unobserved measurements, conditional upon the observed ones. (3) we fit a model to the so-augmented data by using MI techniques to draw values for the unobserved components, conditional upon the observed outcomes and correct pattern-specific density in model (22). As mentioned above, MI consists of three steps: imputation, analysis and combination. The identifying step corresponds to the so-called imputation step, and the final model corresponds to the analysis step. Finally, the combination step, is where the inferences from a number of imputations are drawn together and combined into a single one. After applying each of the three identifying restrictions, model (19) is analyzed again. For MI technique, we again use M=5. Namely, we ended up with totally five multiply-imputed data sets for each choice of identifying restriction strategy which can be analysed, using several possible models. Once the imputations have been generated, the final analysis model from each completed data sets is fitted and MI inference conducted. The parameter and precision estimates are obtained using classical MI machinery. In particular, the asymptotic covariance matrix of the form (5). The analysis of identifying restrictions, fitting of imputed data, and a combination of the results into a single inference was implemented using the SAS macro. This SAS macro dealt with the analysis of the three types of identifying restrictions as follows. First, fit the linear mixed model per pattern using PROCs SORT and MIXED. Second, complete the data using ACMV, CCMV and NCMV restrictions using PROCs IML and MI. Third, analyze the 5 complete data sets using a linear-mixed model using PROC MIXED. Fourth, combine the results from the 5 model fits using PROC MIANALYZE. This SAS macro is available from the authors on request.

4.4.3 Results and discussion from example 2

Table 3 shows the results for the marginal measurement model as well as in the logistic dropout model. Our main interests lie in the marginal treatment effect. There is no overall treatment effect and p-values between the three models do not vary too much. However, the situation is different for the occasion-specific treatment effects considered here. For all weeks, all four p-values

Table 3: Maximum likelihood and multiple imputation for the parameter estimates (standard errors) and p-values, resulting from the SMs and PMMs

Effect	Parameter	Est.(S.e.)	P-value	Est.(S.e.)	P-value	Est.(S.e.)	P-value
Selection models model							
Measurement model							
		MCAR		MAR		MNAR	
intercept$_6$	β_{11}	243.17 (6.74)	< 0.0001	243.17 (6.74)	< 0.0001	243.17 (6.74)	< 0.0001
intercept$_{12}$	β_{21}	244.93 (6.46)	< 0.0001	244.93 (6.45)	< 0.0001	243.98 (6.46)	< 0.0001
intercept$_{20}$	β_{31}	258.92 (6.69)	< 0.0001	258.91 (6.70)	< 0.0001	258.13 (6.70)	< 0.0001
intercept$_{24}$	β_{41}	257.08 (8.02)	< 0.0001	257.08 (8.03)	< 0.0001	256.28 (8.99)	< 0.0001
treatment$_6$	β_{12}	2.36 (8.69)	0.784	2.36 (8.68)	0.788	2.36 (8.69)	0.732
treatment$_{12}$	β_{22}	6.41 (8.32)	0.478	6.41 (8.36)	0.461	6.54 (8.36)	0.441
treatment$_{20}$	β_{32}	-5.77 (8.78)	0.535	-5.77 (8.79)	0.427	-5.83 (8.79)	0.456
treatment$_{24}$	β_{42}	-2.06 (10.70)	0.883	-2.06 (10.74)	0.856	-2.59 (10.79)	0.912
dropout model							
intercept	ψ_0	-1.88(0.11)		-1.73(0.14)		-1.64(0.27)	
$y_{i,j-1}$	ψ_1			-0.20(0.05)		0.04(0.02)	
$y_{i,j}$	ψ_2					-0.16(0.08)	
-2 log-likelihood		3346.4		3329.3		3327.7	
Identifying restrictions							
		ACMV		CCMV		NCMV	
intercept$_6$	β_{11}	243.17 (6.74)	-	243.17 (6.74)	-	243.17 (6.74)	-
intercept$_{12}$	β_{21}	245.44 (7.06)	< 0.0001	245.36 (6.51)	< 0.0001	245.86 (6.55)	< 0.0001
intercept$_{20}$	β_{31}	255.78 (6.71)	< 0.0001	255.88 (6.83)	< 0.0001	257.99 (6.78)	< 0.0001
intercept$_{24}$	β_{41}	256.59 (8.10)	< 0.0001	256.99			
treatment$_6$	β_{12}	2.36 (8.69)	-	2.36 (8.69)	-	2.36 (8.69)	-
treatment$_{12}$	β_{22}	6.23 (8.39)	0.540	6.16 (8.37)	0.539	5.41 (6.45)	0.716
treatment$_{20}$	β_{32}	-5.98 (8.85)	0.484	-5.13 (8.81)	0.475	-6.73 (8.82)	0.290
treatment$_{24}$	β_{42}	-2.18 (11.03)	0.627	-2.12 (11.64)	0.565	-1.87 (10.13)	0.629

Note: MCAR=missing completely at random; MAR=missing at random; MNAR=missing not at random; Est.=parameter estimates; S.e.=standard errors; ACMV=available case missing values; CCMV=complete case missing values; NCMV=neighboring case missing values.

the treatment effects indicate non-significance, whereas for all cases the p-values are certainly highly significance ($p < 0.0001$) for all intercepts. The LRT test statistic for comparing the MAR and MCAR models is 17.1. The corresponding tail probability from $\chi 2$ on 1 degree of freedom is $p < 0.001$ which is significant. This indicates that there is a significant evidence for MAR. In other words, dropout completely at random can be ruled out in the context of the assumed model.

However, great care has to be taken with such a conclusion using only the data under analysis (Diggle and Kenward, 1994). We now focus on factors which influence dropout. In doing so, in the full SMs, the logistic regression for dropout is modeled based on (20). As can be seen in Table 2, the maximum likelihood estimates for $\psi 1$ (0.04) and $\psi 2$ (-0.16) are not necessarily equal, however, their signs are different. This finding is not surprising. It confirms the argument put forward by Molenberghs and Verbeke (2005). They pointed out that since two subsequent measurements are usually positively correlated, the dropout model can depend on the increment, i.e., $yij - yi,j-1$. The full dropout model estimated from the MNAR process is as follows:

$$\text{logit}[p(Di = j \mid yij-1, yij)] = -1.64 - 0.12yi,j-1 - 0.16(yij - yi,j-1). \qquad (23)$$

We now re-parameterize this fitted model in terms of the increment and the sum of the successive measurements to obtain some insight into this fitted model. Therefore, by rewriting equation (20), the fitted dropout model equals

$$\text{logit}[p(Di = j \mid yij-1, yij)] = \vartheta 0 + \vartheta 1(yi,j + yi,j-1) + \vartheta 2(yij - yi,j-1), \qquad j = 2, 3, 4, 5. \quad (24)$$

Here, $\vartheta 1 = (\psi 1 + \psi 2)/2$ and $\vartheta 2 = (\psi 1 - \psi 2)/2$, representing dependence on level and increment in the serum cholesterol, and these quantities are likely to be much less strongly correlated than are yij and $yi,j-1$. Thus from the fitted MNAR model in equation (24), we have

$$\text{logit}[p(Di = j \mid yij-1, yij)] = -1.64 - 0.06(yi,j + yi,j-1) + 0.10(yij - yi,j-1), \quad (25)$$

which is to say that the probability of dropout increases with larger negative increments. In other words, those patients with a greater increase in the overall level of the serum cholesterol from the previous week have a higher probability of dropping out of the experiment.

The results in Table 3 show that the association p-values for the marginal effect assessments are all nonsignificant, their p-values being all greater than 0.05. However, the association p-values for the intercepts are highly significant ($p < 0.0001$), in line with the p-values obtained from the SMs analysis. These findings confirm those obtained from the SMs formulation which gives more weight to this conclusion. Overall, it is clear that there is strong evidence for no significant treatment in the context of serum cholesterol data. This explains the fact that PMMs using identifying restrictions strategy play a very similar role to the modelling assumptions in the SMs based on Diggle-Kenward type (Michiels et al., 1999). Therefore, one can put more confidence in this conclusions as many authors (for example, Michiels et al., 1999) have argued that greater confidence in a conclusion can be reached when the analysis of joint applications of these models leads to essentially similar inference when assessing significant effects, such as marginal treatment effects. In contrast to SMs, the use of CCMV, NCMV and ACMV restrictions strategy did not allow an estimation of whether the dropout process is MNAR or not, because of differences in the modelling assumptions. According to Molenberghs et al. (1998), the identifying restrictions in PMMs context can be used only to relate the model to a MAR mechanism. Thus an important issue is to equate results for both the ACMV and MAR to make a clear and useful connection between the selection model and the pattern mixture model framework (Verbeke and Molenberghs, 2000; Kenward et al., 2003). With this in mind, the same is true for the selection model, MAR-based ACMV restrictions indicating non-significant treatment effects at all weeks. This can be explained to mean that the treatment effect appear to be independent of the ACMV (MAR) assumption. Although corresponding models include the same effects, the estimates for ACMV are slightly different to those for

MAR. These slight differences are to be expected as argued in Kenward et al. (2003) that both models are similar in spirit but not necessarily identical.

5 Conclusion

This study reviewed some of the key modelling strategies and basic issues in statistical data analysis to address the dropout problem in longitudinal clinical trials. The main objective was to provide an overview of issues and different methodologies in the case of dropout in longitudinal clinical trials due to a result of patients dropping out of a study. The focus was on dropout with a monotone pattern. The methodologies investigated for handling dropout were: Deletion methods, Imputation-based methods, Data augmentation methods, Likelihood-based methods, and MNAR-based methods. For each methodology, several methods that are commonly used to deal with dropout are presented. Two application examples were presented to highlight two ways. On the one hand, the first application example served to demonstrate comparison of existing approaches providing useful and important information regarding their applications. The second application example, on the other hand, provided techniques that might serve as tools in the context of a sensitivity analysis thereby broadening the possibilities under such.

In the context of planning data collection, clinical trial study designers must think of study or clinical designs and data collection strategies that minimize dropouts since data collection plays an important role in the problem of dropout for a specific study. This means that careful planning can reduce the amount of dropouts although there is no rule concerning the level of dropouts that can be acceptable. Thus, at the analysis stage, how to handle the dropout and how to minimize the amount of dropouts are main issues that must be considered when planning and designing a study for data collection. In the presence of dropout, knowing the reasons why the data were dropping out, as well as exploring the dropout pattern become very important and helpful in choosing the right statistical procedures to approximate dropouts. In fact, there is no universal technique for handling all dropout situations, however, there are some rules that can be considered. As such, it is necessary to design a study where the potential pattern of dropouts is considered when specifying the primary analysis. In conclusion, it is important to clearly understand the limitations of the different techniques for handling dropout, and the current study however attempted to review many of these limitations.

REFERENCES

[1]. Allison, P. D. (2002). Missing data. Thousand Oaks, CA: Sage.

[2]. Baker, S. G. (1992). A simple method for computing the observed information matrix when using the EM algorithm with categorical data. Journal of Computational and Graphical Statistics, 1, 63-76.

[3]. Beunckens, C., Sotto, C., and Molenberghs, G. (2008). A simulation study comparing weighted estimating equations with multiple imputation based estimating equations for longitudinal binary data.

[4]. Breslow, N. E., and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. Journal of American Statistical Association, 88, 9-25.

[5]. Brown, C. H. (1983). Asymptotic comparison of missing data procedures for estimating factor loadings. Psychometrika, 48, 269-291. Computational Statistics and Data Analysis, 52, 1533-1548.

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

[7]. Cook, R. D. (1986). Assessment of local influence (with discussion). Journal of the Royal Statistical Society, Series B, 48, 133-169.

[8]. Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). Journal of Royal Statistical Society: Series B, 39, 1-38.

[9]. Diggle, P. J. and Kenward, M. (1994). Informative dropout in longitudinal data analysis (with discussion). Applied Statistics, 43, 49-94.

[10]. Dmitrienko, A., Offen, W. W., Faries, D., Chuang-Stein, C., and Molenberghs G. (2005). Analysis of clinical trial data using the SAS system, Cary, NC: SAS Publishing.

[11]. Forster, J. J. and Smith, P. W. (1998). Model-based inference for categorical survey data subject to nonignorable non-response. Journal of the Royal Statistical Society, Series B, 60, 57-70.

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

[13]. Kenward, M. G. (1998). Selection models for repeated measurements with nonrandom dropout: An illustration of sensitivity. Statistics in Medicine, 17, 2723-2732.

[14]. Kenward, M., Molenberghs, G., and Thijs, H. (2003). Pattern mixture models with proper time dependence. Biometrika, 90, 53-71.

[15]. Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. Biometrics, 38, 963-974.

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

[17]. Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. Journal of the American Statistical Association, 88, 125-134.

[18]. Little, R. J. A. (1994). A class of pattern-mixture models for normal incomplete data. Biometrika, 81, 471-483.

[19]. Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. Journal of the American Statistical Association, 90, 1112-1121.

[20]. Little, R. J. A. and Rubin, D. B. (1987). Statistical analysis with missing data. New York: John Wiley.

[21]. Little, R. J. A. and Rubin, D. B. (2002). Statistical analysis with missing data (2nd edn). New York: John Wiley and Sons.

[22]. Little, R. J., DAgostino, R., Cohen, M. L., Dickersin, K., Emerson, S. S., Farrar, J., Frangakis, C., Hogan, J. W., Molenberghs, G., Murphy, S. A., Neaton, J. D., Rotnitzky, A., Scharfstein, D., Shih, W, J., Siegel, J. P., and Stern, H. (2012). The prevention and treatment of missing data in clinical trials. The New England Journal of Medicine, 367, 1355-1360.

[23]. Louis, T. A. (1982). Finding the observed information when using the EM algorithm. Journal of Royal Statistical Society: Series B, 44, 226-233.

[24]. Michiels, B., Molenberghs, G., and Lipsitz, S. R. (1999). Selection models and pattern mixture models for incomplete data with covariates. Biometrics, 55, 978-983.

[25]. Milliken, G. A. and Johnson, D. E. (2009). Analysis of messy data. Design Experiments, volume 1. Second Edition. Chapman and Hall/CRC.

[26]. McKnight, P. E., McKnight, K. M., Sidani, S., and Figueredo, A. J. (2007). Missing data: A gentle introduction. The Guilford Press: New York.

[27]. Molenberghs, G., Michiels, B., Kenward, M., and Diggle, P. J. (1998). Missing data mechanism and pattern mixture models. Statistica Neerlandica, 52, 135-161.

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

[29]. Molenberghs, G. and Verbeke, G. (2005). Models for discrete longitudinal data. New York: Springer.

[30]. Molenberghs, G. and Kenward, M. G. (2007). Missing data in clinical studies. England: John Wiley and Sons.

[31]. Molenberghs, G., Beunckens, C., Sotto, C., and Kenward, M. (2008). Every missing not at random model has got a missing at random counterpart with equal fit. Journal of Royal Statistical Soceity: Series B, 70, 371-388.

[32]. Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. Journal of the American Statistical Association, 89, 846-866.

[33]. 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 American Statistical Association, 90, 106-121.

[34]. Robins, J. M., Rotnitzky, A. and Scharfstein, D. O. (1998). Semiparametric regression for repeated outcomes with non-ignorable non-response. Journal of the American Statistical Association, 93, 1321-1339.

[35]. Rubin, D. B. (1976). Inference and missing data. Biometrika, 63, 581-592.

[36]. Rubin, D. B. (1987). Multiple imputation for non-response in surveys. New York: John Wiley.

[37]. Rubin, D. B., Stern, H. S., and Vehovar, V. (1995). Handling dont know survey responses: the case of the Solvenian plebiscite. Journal of the American Statistical Association, 90, 822-828.

[38]. Satty, A. and Mwambi, H. (2012). Imputation methods for estimating regression parameters under a monotone missing covariate pattern: A comparative analysis. South African Statistical Journal, 46, 327-356

[39]. Schafer, J. L. (1997). Analysis of incomplete multivariate data. New York: Champan and Hall.

[40]. Schafer, J. L. and Graham, J. W. (2002). Missing data: Our view of the state of the art. Psychological Methods, 7, 147-177.

[41]. Schafer, J. L. (2003). Multiple imputation in multivariate problems when the imputation and analysis models differ. Statistica Neerlandica, 57, 19-35.

[42]. Schoenfield, L. J. and Lachin, J. M. (1981). The steering committee, and the NCGS group, Chenodiol (Chenadeoxycholic Acid) for Dissolution of Gallstones: The National Cooperative Gall- stone Study. Annals of Internal Medicine, 95, 257-282.

[43]. Verbeke, G. and Molenberghs, G. (2000). Linear mixed models for longitudinal data. New York: Springer.

[44]. Verbeke, G., Lesaffre, E., and Spiessens, B. (2001). The practical use of different strategies to handle dropout in longitudinal studies. Drug Information Journal, 35, 419-434.

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

[46]. 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, 283-297.

[47]. Zhao, L. P. and Lipsitz, S. (1992). Design and analysis of two-stage studies. Statistics in Medicine, 11, 769-782.