Handling Dropouts in Longitudinal Clinical Trials Aata: Likelihood-based Analysis Versus Inverse Probability Weighting

[1]**Ali Satty** and [2]**Ali Basher Abdullah Babikir**
[1]University of Kwazulu Natal, Durban, South Africa
[2]School of Statistics and Actuarial Science
Elnilain University, Khartoum, Sudan
E-mail: alisatty1981@gmail.com

ID# (2754)
Rec.: 13/06/2013
In-revised: 04/08/2013
Corresponding Author;
Ali Satty E: mail:
alisatty1981@gmail.com

KEYWORDS

Incomplete longitudinal data; Likelihood-based analysis; Inverse probability weighted (IPW); Dropout; Missing at random (MAR); Linear mixed model (LMM).

ABSTRACT

Dropout is a pervasive problem in longitudinal clinical trials, and it is the result mainly of non-responses due to individuals who leave the study and are therefore lost to follow-up. The current paper deals with incomplete longitudinal clinical trials data when there are dropout. Statistical methods that ignore the mechanism for dropouts are susceptible to biased inference. This article focuses on dropouts missing at random (MAR). The study demonstrates application and the performance of likelihood-based and inverse probability weighting (IPW) in handling dropout in longitudinal continuous responses. The main objective of this paper is to compare the performance of these methods under different dropout rates. Data from a study with individual heart rate as the outcome is used to investigate the performance of the considered methods. Based on this longitudinal clinical trial data, results from IPW will be compared with those obtained from likelihood-based analysis. The performance of these methods are compared in terms of bias and efficiency.

تحليل إحصائي لمقارنة طريقة معكوس الاحتمال المُرَجَح وطريقة الإمكان الأعظم لمُعالجة البيانات الطبية المفقودة عشوائياً

[1]**علي ساتي** ، و[2]**علي بشير عبد الله بابكر**
[1]جامعة كوازولو ناتال ، ديربان، جنوب افريقيا
[2]مدرسة الإحصاء والعلوم، جامعة النيلين، الخرطوم، السودان
بريد الكتروني:alisatty1981@gmail.com

رقم المسودة: # (2754)
إستلام المسودة: 2013/06/13
إستلام المُعَدَلة: 2013/08/04
الباحث المُراسل: علي ساتي
بريد إلكتروني:
alisatty1981@gmail.com

الكلمات الدالة

طريقة معكوس الاحتمال المرجح، بيانات التجارب السريريه الطوليه، طريقة الامكان الاعظم، البيانات المفقوده عشوائياً

المُستلخص

تعتبر مشكلة البيانات المفقوده واحده من اهم مشاكل بيانات التجارب في المعالجات السريريه طويلة المدى. يهتم هذا البحث بدراسة طريقتين لمعالجة هكذا البيانات المفقوده عشوائيا. الطريقه الاولى هي المعروفة بطريقة معكوس الاحتمال المرجح اما الطريقه الثانيه فتعرف بطريقة الامكان الاعظم. الهدف الرئيسي من هذه الورقه هو مقارنة أداء الطريقتين على ضوء نسب مختلفه للبيانات المفقوده. لتحقيق هذا الهدف يتم استخدام بيانات تجارب سريريه خاصه بمرضي القلب. و تتم مقارنة نتائج الطريقتين باستخدام بعض المقاييس الاحصائيه منها التحيز و الكفاءة.

Introduction

Longitudinal clinical trials studies are frequently designed to collect data on every patient within a sample at each measurement occasion. In such studies, dropout often arises, in the sense that an individual's outcome can be missing at one follow-up time before the end of the follow-up period for whatever reason, resulting in a monotone missingness pattern. This paper however only pays attention to the monotone missing data pattern that results from attrition. In this pattern, when patient drops out from the study prematurely, no more measurements are obtained on that patient. The implication of dropout is best understood by considering the process (i.e., the mechanisms) leading to the incompleteness. Based on classifications given by Rubin (1976, 1987), these mechanisms can be classified into three categories: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR).

Methods and Materials

There are two common methods that can be used for dealing with longitudinal clinical trials data when there are dropouts missing at random (MAR). One is the so-called inverse probability weighted (IPW) estimating equations. In this method, the complete measurements are weighted using the inverse of their probabilities of being observed in order to adjust for dropout. IPW comes from Robins *et al*, (1995). The authors stated that IPW deals with incomplete longitudinal data arising from an MAR mechanism. In the context of survey analysis, IPW was first discussed in Horvitz and Thompson (1952). The method 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 (Grace and Wenqing, 2009). More details of this technique can be found in Fitzmaurice et al, (1995), Yi and Cook (2002a, 2002b), Carpenter *et al,* (2006), and Seaman and White (2011). However, Robins *et al,* (1995), Robins and Rotnitzky (1995) and Scharfstein *et al,* (1999) have proposed improved IPW estimates that are theoretically more efficient. However, such estimates require the dropout mechanism to be MAR.

An alternative technique that is valid under the MAR assumption is likelihood-based analysis. In contrast to IPW, this method uses the observed measurement without weight. The strength of the likelihood-based analysis relies on the accurate formulation of the likelihood of the data as it is. In doing so, under valid MAR mechanism, suitable adjustments can be made to parameters at times when data are prone to incompleteness due to the within-patient correlation. For incomplete longitudinal clinical trials data, a mixed model only needs a dropouts are MAR (Mallinckrodt *et al, ,* 2003a, 2003b; Verbeke and Molenberghs, 2000). These mixed effects models permit the inclusion of patients with dropouts at some time points, including both dropout (which is a special case of monotone missingness pattern) and intermittent missingness patterns (Verbeke and Molenberghs, 2000). For continuous measurement data, this amounts to the general linear mixed model (Verbeke and Molenberghs, 2000).

The current study focuses on the comparison of LMM and IPW to handling incomplete longitudinal clinical trials data due to dropout. The application will be confined to the continuous measurements. Our goal will be to compare the performance of these two methods, under three different dropout rates. In both methods the dropout mechanism is assumed to be MAR. In order to compare the performances of the two methods, the data set we use is originally complete (no dropouts), and the dropout will be created by generating missing measurements at random. The comparison will be based on a heart rate trial data, which gives heart rate observations for patients exposed to three different treatments as reported in Milliken and Johnson (2009). The performance of these two approaches will be assessed on two criteria, namely bias and efficiency. In Section 1, we present the notation and concepts of possible mechanisms that can lead to the dropout process. In Section 2, the two approaches mentioned above (Inverse Probability Weighting (IPW) & Likelihood-based LMM) are then considered in more detail as the principle approaches to be used in the analysis. Section 3, contains the design of the application study and offers a description of the dataset used in the analysis in detail (description of the data, model formulation, generating dropouts and the MAR mechanism, analysis of IPW, and assessment criteria) . The results of the study based on the generated dropout data in our application , and discussion of the results considered as conclusion of the study.

(1) Dropout Mechanism

Suppose that N individuals are to be observed at n occasions. For the i th individual $(i = 1,2,...N)$ we can form a $(n \times 1)$ vector $Y_i = (Y_{i1},...,Y_{in})'$, where Y_{ij} is the j th outcome for individual i, which can be continuous or discrete depending on the study problem. Each individual has a $(n \times p)$ covariate matrix Y_i. The

covariates may be both time stationary and time varying. In longitudinal studies, individuals can be unobserved at all n occasions on account of some stochastic missing data mechanism. Now, suppose R_i is a $(n\times 1)$ random vector for the i th individual, whose j th component R_{ij} equals 1 when Y_{ij} is fully observed, and equals 0, if not. The full data information for the i th individual are given jointly by Y_i and R_i, with a joint distribution that can be expressed as

$$f_{y,r} = (y_i, r_i \mid X_i, \theta, \gamma) = f_r(r_i \mid y_i, X_i, \gamma) f_y(y_i \mid X_i, \theta), \quad \textbf{(1)}$$

where θ and γ are vectors that parameterize the joint distribution. The "missing data mechanism", $f_r(r_i \mid y_i, X_i, \gamma)$ is parameterized by γ . In general, the mechanism of missing data can depend on the full vector of responses, Y_i (including possibly unobserved component of Y_i) and the matrix of covariates X_i. We denote the observed and unobserved components of Y_i by Y_i^o and Y_i^m, respectively. Rubin (1976, 1987), specified three distinct missing data mechanisms. First, data that is MCAR, meaning that the missingness process does not depend on Y_i, i.e., $P(R_i \mid Y_i^o, Y_i^m, X_i, \gamma) = P(R_i \mid X_i, \gamma)$. Second, the missing data is said to be MAR if, the missingness process depends on the observed responses and probably on measured covariates, but not on the unobserved responses, i.e., $P(R_i \mid Y_i^o, Y_i^m, X_i, \gamma) = P(R_i \mid Y_i^o, X_i, \gamma)$. The third missing data mechanism allows the missingness process to depend on the unobserved responses, and here such a process is called missing not at random (MNAR). However, an MNAR process is also allowed to depend on the observed outcomes, or in probability terms, $P(R_i \mid Y_i^o, Y_i^m, X_i, \gamma) = P(R_i \mid Y_i^m, Y_i^o, X_i, \gamma)$. In terms of likelihood based inference, Rubin (1987) showed that an MCAR is a special case of MAR, and these two mechanisms are referred to as being "ignorable". In contrast, an MNAR mechanism is referred to as a "non-ignorable" mechanism. More specifically, the likelihood inference is based on

$$L(\theta, \gamma \mid X_i, y_i, r_i) \propto f(y_i^o, r_i \mid X_i, \theta, \gamma) = f(y_i^o, r_i \mid \theta, \gamma) = \int f(y_i, r_i \mid X_i, \theta, \gamma) dy_i^m. \quad \textbf{(2)}$$

Therefore,

$$f(y_i^o, r_i \mid \theta, \gamma) = \int f(y_i^o, y_i^m \mid X_i, \theta) f(r_i \mid y_i^o, y_i^m, X_i, \gamma) dy_i^m. \quad \textbf{(3)}$$

Under dropout MAR process, the likelihood contributions factor is:

$$f(y_i^o, r_i \mid \theta, \gamma) = \int f(y_i^o, y_i^m \mid X_i, \theta) f(r_i \mid y_i^o, X_i, \gamma) dy_i^m = f(y_i^o \mid X_i, \theta) f(r_i \mid y_i^o, X_i, \gamma). \quad \textbf{(4)}$$

The likelihood in (4) factorizes into two components of the same functional form as the general factorization of the full data (Y_i, R_i) given in (1). Further, if the parameters θ and γ are disjoint which is to say the parameter space of the full vector vector $(\theta', \gamma')'$ is the product of the individual parameter spaces, the so-called separability condition, then inference can be based on the marginal observed data density only. Hence, when the separability condition is satisfied via a likelihood framework, ignorability is equivalent to MAR and MCAR. However, an MNAR mechanism is defined as a "non-ignorable" mechanism in the context of the likelihood framework. See, Little and Rubin (2002) for details on the derivation of the contribution to the likelihood attributable to the missingness mechanisms. Recall that our focus in this study is on missing data due to patient dropout, that is; all components of Y_i will be missing, and all components of R_i will be 0 starting from the dropout time. The dropout time for the i th individual can be defined by introducing a quantitative variable

$$D_i = 1 + \sum_{t=1}^{n} R_{it}, \quad \textbf{(5)}$$

and hence the model for dropout process can be rewritten as

$$\iota_{id_i} = f(r_i \mid y_i, X_i, \gamma) = Pr(D_i = d_i \mid y_i, X_i, \gamma). \quad \textbf{(6)}$$

where d_i is a realization of the the variable D_i. In Equation (5), it is assumed that all subjects are observed on the first occasion so that D_i takes values between 2 and $n+1$. The maximum value $(n+1)$ corresponds to a complete measurement sequence. Using Equation (6), a dropout missing completely at random (MCAR) model reduces to $P(D_i = d_i \mid Y_i, X_i, \gamma) = P(D_i = d_i \mid X_i, \gamma)$, while the dropout missing at random (MAR) model is given by: $P(D_i = d_i \mid Y_i, X_i, \gamma) = P(D_i = d_i \mid Y_i^o, X_i, \gamma)$, where dependence on Y_i is only through Y_i^o.

(2) Methods for Handling Dropouts

(2.1) Inverse Probability Weighting (IPW)

This method is a standard method used for handling MAR dropout in longitudinal clinical trials (Robins *et al*, 1995), however it requires specification of a dropout model in terms of observed measurements and/or covariates. Generally, IPW is used in marginal models for discrete measurements than for continuous measurements. In this study, IPW is adopted for dealing with continuous measurements in order to correct the bias that is caused by MAR dropout. The key idea behind IPW is that if individual i has a probability of being observed at occasion t of λ_{it}, then, this individual should be given weight, ω_{ij}, so as to minimize the bias caused by dropouts in the analysis. The weight ω_{ij}

for the i th individual at time j is assigned as inverse of the cumulative product of fitted probabilities, $\hat{w}_{ij}(\hat{\alpha})=(\hat{\lambda}_{i1}(\hat{\alpha})\times\hat{\lambda}_{i2}(\hat{\alpha})\times...\times\hat{\lambda}_{ij}(\hat{\alpha}))^{-1}$, where α is a $(q\times 1)$ vector of unknown parameters. In order to discuss the idea of what weights are, we follow illustration provided by Carpenter *et al,* (2006). Suppose that we have the following data, then the average response is 3.

Group	A	B	C
Response	222	333	444

However if we have missing values as shown in the table below, then the average response is 19/6, which is biased.

Group	A	B	C
Response	2??	333	?44

In order to correct this bias, we calculate the probabilities of being observed in each group corresponding to 1/3 in group A, 1 in group B and 2/3 in group C. We thereafter calculate a weighted average, where each observation is weighted by 1/ [Probability of observed response]. In this case the weighted average is given by

$$\frac{2\times\frac{3}{1}+(3+3+3)\times 1+(4+4)\times\frac{3}{2}}{\frac{3}{1}+1+1+1+\frac{3}{2}+\frac{3}{2}}=3 \qquad \textbf{(7)}$$

which now corrects the bias. The conclusion to be drawn from this simple illustration is that IPW has eliminated the bias, by reconstructing the full population by up-weighting the data from individuals who have small chance of being observed. Generally, it may give biased, but consistent, parameter estimates (Carpenter *et al*, 2006). To discuss the above mentioned idea of IPW in longitudinal data setting, we now describe the IPW approach, thereby illustrating how IPW can be incorporated into the conventional generalized estimating equation (GEE) by Liang and Zeger (1986), as based on the article by Robins *et al,* (1995). The primary idea behind GEE methodology is to generalize the usual univariate likelihood equations by introducing the covariance matrix of the vector of response, Y_i. The GEE methodology is used to model the marginal expectation of responses as a function of a set of of covariates. We briefly introduce the classical form of GEE (Liang and Zeger, 1986). Let $X_i=(x_{i1},...,x_{in})'$ denote an $(n\times p)$ covariates matrix, where its t th row is given $x'_{it}=(x_{it1},...,x_{itp})'$ based on p predictor variables or covariates, Y_{it} denote the response variable and hence $y_i=(y_{i1},...,y_{in})'$ the $(1\times n)$ observed response vector, and $\mu_{it}=E(y_{it}),\ i=1,...,N$ and $t=1,...,n$. Now, assume the marginal regression model is given as

$$g(\mu_{it})=x'_{it}\beta, \qquad \textbf{(8)}$$

where β is the $(p\times 1)$ regression parameters of interest and $g(.)$ is a link function, a function of the mean response. Assume the $(n\times n)$ covariance matrix for Y_i is: $\Sigma(\varphi)=\phi A_i^{1/2}R(\rho)A_i^{1/2}$, where A_i is a diagonal matrix of variance functions, $R(\rho)$ is the working correlation matrix of Y_i as a function of ρ the correlation parameter, and ϕ is the dispersion parameter. The collection of parameters in the covariance matrix are assumed to contained in the parameter vector φ. Then the GEE estimators for regression parameters are the solutions of

$$\sum_{i=1}^{n}D_i\Sigma(\varphi)^{-1}(Y_i-\mu_i)=0, \qquad \textbf{(9)}$$

where $D_i=\dfrac{\partial\mu_i}{\partial\beta}$ is the derivative matrix of the mean vector μ_i with respect to β. The GEE methodology is very popular especially for analysis of marginal models for discrete responses than for continuous responses. However, in this study, we restrict our attention to the continuous response. Consequently, the following assumptions can be made for the marginal models with the continuous response, Y_{it}

(2.1.1) The mean of Y_{it} is related to the covariates by an identity link function: $\mu_{it}=\eta_{it}=x'_{it}\beta$ The link function $g(.)$ generally relates the expected values, μ_i of the response vector, Y_i to the covariate matrix X_i. It takes the general form $g(\mu_i)=\eta_i=X_i\beta$, where η_i denotes the linear predictor vector whose t th row is $g(\mu_{it})=\beta_1x_{it1}+\beta_2x_{it2}+\cdots+\beta_px_{itp}$This function, i.e.,$g(.)$, should be monotone and differentiable. Thus, in the case of monotonocity, we can define the inverse function $g(.)^{-1}$ by the relation $g^{-1}(g(\mu_i))=\mu_i$. Here, we note that the choice of link function depends on the distributional assumptions on the data. Therefore, for a continuous response with normal assumption, as in our case, the link function is an identity link: $g(\mu_i)=\mu_i$ and the inverse simply $\mu_i=g(\mu_i)$ Under this identity link, the expected value of the response is simply a linear function of the covariates multiplied by their regression coefficients.

(2.1.2) The variance of each Y_{it}, conditional on the effects of the covariates, is ϕ and does not depend on the mean response. Namely,

$v(\mu_{ij}) = 1$ is a known "variance function", therefore $Var(Y_{it}) = v(\mu_{it})\phi = \phi$. Here ϕ denotes the variance of the conditional normal distribution of the response, given the covariates. The assumption that the variance is constant over time may be unrealistic and to relax it, a separate scale parameter, ϕ_t could be estimated at the t th occasion if the longitudinal design is balanced on time.

(2.1.3) The within-individual correlation among repeated responses is modelled by assuming, for example, a first-order autoregressive AR(1) covariance structure, $Corr(Y_{ij}, Y_{ik}) = \rho^{|k-j|}$, which stand for the pairwise correlation between observations, for all j and k and $0 \le \rho \le 1$. The AR(1) specifies homogeneous variances. In addition, it specifies that the correlations between observations on the same subject are not equal, but decrease toward zero with increasing length of the time interval between observations. In the context of the marginal models with an identity link function, the generalized least square means (GLS) of β can be considered as a special case of the GEE. Therefore, the estimates of parameters in marginal model for continuous response with an identity link are

$$\hat{\beta} = \{\sum_{i=1}^{n} X_i' \hat{\Sigma}_i^{-1} X_i\}^{-1} \sum_{i=1}^{n} (X_i' \hat{\Sigma}_i^{-1} Y_i), \quad \textbf{(10)}$$

where $\hat{\Sigma}_i$ is the REML (Restricted Maximum Likelihood Estimation, that can be used to find the best unbiased estimates (Verbeke and Molenberghs, 2000)) estimate of Σ_i and

$$Cov(\hat{\beta}) = \{\sum_{i=1}^{n} X_i' \hat{\Sigma}_i^{-1} X_i\}^{-1} \sum_{i=1}^{n} (X_i' \hat{\Sigma}_i^{-1} \hat{V}_i \hat{\Sigma}_i^{-1} X_i) \{\sum_{i=1}^{n} X_i' \hat{\Sigma}_i^{-1} X_i\}^{-1}, \quad \textbf{(11)}$$

where $\hat{V}_i = (Y_i - X_i\beta)(Y_i - X_i\beta)$ is an estimate of $Var(Y_i)$ which yields a robust estimator of $Cov(\hat{\beta})$ when substituted in equation (12). With incomplete data that are MAR, the GEE method provides inconsistent estimates of the model parameters (Liang and Zeger, 1986). In weighted generalized estimating equation (WGEE), an individual's contribution to the usual GEE is weighted by the inverse probability of dropout at particular time point, given the individual did not leave or dropout in any of the periods occasions (Robins *et al*, 1995). Therefore, setting all assumptions that are made in this section together, we will get valid parameter estimates in longitudinal studies with MAR dropout by solving the weighted estimating equations

$$\sum_{i=1}^{n} (Y_i - X_i\beta)' \Sigma_i^{-1} W_i(\hat{\alpha}) (Y_i - X_i\beta) = 0, \quad \textbf{(12)}$$

where $W_i(\hat{\alpha})$ is a diagonal matrix which contains inverse probability weights for i th patient, $W_i(\hat{\alpha}) = diag\{\hat{w}_{i1}(\hat{\alpha}), ..., \hat{w}_{in_i}(\hat{\alpha})\}$ for $j = 2, ..., n_i$, $\hat{w}_{i1} = 1$, and $\Sigma_i = A_i^{1/2} R(\rho) A_i^{1/2}$ is a $n \times n$ working covariance matrix for Y_i and $R(\rho)$ is a $n \times n$ working correlation matrix, which are assumed known. The missingness is taken into account through specification of a $(n \times n)$ diagonal weighting matrix of $W_i(\hat{\alpha})$ $W_i(\hat{\alpha}) = diag(R_{i1}\hat{w}_{i1}(\hat{\alpha}), ..., R_{in}\hat{w}_{in}(\hat{\alpha}))$ and $R_{it} = 1$ if the i th subject is observed at time t, and 0 for the unobserved time. The weight, w_{ij} is the inverse of the probability that the i th subject is observed at the j th time, which is often unknown and needs to be estimated. It requires modeling the missing process in order to obtain the weights w_{ij}. We denote $\lambda_{ij}(\alpha) = P(R_{ij} = 1 \mid R_{i(j-1)} = 1, X_i, Y_i, \alpha)$ as the probability of a response being observed at time j for the i th subject given the subject is observed at the time $j-1$. If the missingness is assumed to be MAR, we have

$$\lambda_{ij}(\alpha) = P(R_{ij} = 1 \mid R_{i(j-1)} = 1, X_i, Y_{i1}, ..., Y_{i(j-1)}, \alpha), \quad \textbf{(13)}$$

where the missingness mechanism only depends on observed data and may be specified up to a $(q \times 1)$ vector of unknown parameters, α. Here, λ_{ij} can be modeled as a logistic regression model of Z_{ij}, a vector of predictor, which may include missingness indicator variables, covariates and previous responses:

$$logit\lambda_{ij}(\alpha) = Z_{ij}'\alpha, \quad \textbf{(14)}$$

or by inverting the logit function we have:

$$\lambda_{ij}(\alpha) = \frac{e^{Z_{ij}'\alpha}}{1 + e^{Z_{ij}'\alpha}}. \quad \textbf{(15)}$$

The log partial likelihood for i th subject takes the form

$$\ell(\alpha) = \sum_{j=1}^{n} \sum_{i=2}^{j_i} R_{i(j-1)} log\{\lambda_{ij}(\alpha)^{R_{ij}} [1 - \lambda_{ij}(\alpha)]^{1-R_{ij}}\} \quad \textbf{(16)}$$

Differentiation of (16) in terms of α gives the estimating equations

$$S_i(\alpha) = \{\sum_{i=1}^{N} \sum_{i=2}^{j_i} R_{i(j-1)} [R_{ij} - \lambda_{ij}(\alpha)]\} \quad \textbf{(17)}$$

Setting (17) equal to zero yields $\hat{\alpha}$, therefore, we can obtain estimate of $\lambda_{ij}(\alpha)$, which is $\hat{\lambda}_{ij}(\hat{\alpha})$. According to Hogan *et al* (2004), in addition to MAR dropout, two assumptions must be fulfilled, to provide consistent estimates of parameters β in weighted method. *First assumption* (Non-zero probability of remaining in study): Conditionally on past history of observed responses and covariates, the probability that individual i is still in the study at time j is bounded away from zero or formally, $p[R_{ij}=1 \mid R_{i(j-1)}=1, X_i, Y_i] > \delta > 0$. *Second assumption* (Correct specification of dropout model): The probability of dropout model must be correctly specified, i.e., $\lambda_{ij}(\alpha) = p[R_{ij}=0 \mid R_{i(j-1)}=1, X_i, Y_{i(j-1)}]$. When MAR and monotone missingness assumptions hold, the probabilities of remaining in the study, $\pi_{ij}(\alpha) = p[R_{ij}=1 \mid R_{i(j-1)}=1, X_i, Y_{i1},...,Y_{i(j-1)}, \alpha]$

$$= \prod_{k=1}^{j} \{1-\lambda_{ik}(\alpha)\}. \qquad \textbf{(18)}$$

Thus, the weight $\hat{w}_{ij}(\hat{\alpha})$, the inverse of the unconditional probability of being observed at time j, can be calculated as,

$$\hat{w}_{ij}(\hat{\alpha}) = \frac{1}{1\times(1-\hat{\lambda}_{i1}(\hat{\alpha}))\times...\times(1-\hat{\lambda}_{ij}(\hat{\alpha}))}, \; i=2,...,J, \qquad \textbf{(19)}$$

and $\hat{w}_{ij}(\hat{\alpha}) = 1$ for $j=1$. Therefore, if the above two assumptions due to Hogan *et al,* (2004) hold, and if dropout occurs according to the MAR mechanism, then the estimators of the parameters $\hat{\beta}$ in the weighted marginal model for a continuous response with an identity link will be of the form

$$\hat{\beta} = \{\sum_{i=1}^{n} X_i' \hat{\Sigma}_i^{-1} W_i(\hat{\alpha}) X_i\}^{-1} \sum_{i=1}^{n} (X_i' \hat{\Sigma}_i^{-1} W_i(\hat{\alpha}) Y_i), \qquad \textbf{(20)}$$

and,

$$Cov(\hat{\beta}) = \{\sum_{i=1}^{n} X_i' \Sigma_i^{-1} W_i(\hat{\alpha}) X_i\}^{-1} \qquad \textbf{(21)}$$

$$(\sum_{i=1}^{n} X_i' \Sigma_i^{-1} W_i(\hat{\alpha}) W_i(\hat{\alpha})' X_i)\{\sum_{i=1}^{n} X_i' \Sigma_i^{-1} W_i(\hat{\alpha}) X_i\}^{-1},$$

where $\hat{\beta}$ is consistent for β, and $\hat{\alpha}$ is a consistent estimator of α under a correctly specified model, $\lambda_{ij}(\alpha)$.

(2.2) Likelihood Based MAR

The second method for handling incomplete longitudinal clinical trials data is the likelihood-based method of using available data instead of weight. For continuous measurements, Laird and Ware (1982) proposed the Likelihood-based mixed effects models which are valid under MAR dropout. This likelihood-based MAR analysis is also termed likelihood-based ignorable analysis, or direct likelihood analysis (Verbeke and Molenberghs, 2005). In contrast to IPW, Likelihood-based analysis uses the observed data without the need of weight. Namely, no additional data manipulation is necessary when a likelihood-based analysis is envisaged, provided the software tool used for analysis is able to handle measurement sequences of unequal length (Molenberghs and Kenward, 2007). In doing so, for the MAR assumption, suitable adjustments can be made to parameters at times when data are prone to incompleteness due to the within-subject correlation. Thus, even when interest lies in a comparison between two treatment groups at the last measurement time, such a likelihood analysis can be conducted without problems since the fitted model can be used as the basis for inference. When a MAR mechanism holds, a direct likelihood analysis can be obtained with no need for modeling the dropout process. It is preferred over the conventional simple methods, particularly when tools like the generalized linear mixed mixed effect models (Molenberghs and Verbeke, 2005) are assumed. The advantage of this method is its simplicity, it can be fitted in standard statistical software without involving additional programming, using such tools as SAS software, PROCs MIXED, GLIMMIX and NLMIXED. The use of these procedures have been illustrated by Verbeke and Molenberghs (2000) and Molenberghs and Verbeke (2005). Despite the flexibility and ease of implementation of likelihood-based method, there are fundamental issues when selecting a model and assessing its fit to the observed data which do not occur with complete data. The method is sensible under linear mixed models in combination with the assumption of ignorability. Such an approach, tailored to the needs of clinical trials, has been proposed by Mallinckrodt *et al* (2001a, 2001b). For incomplete longitudinal clinical trials data, a mixed model only needs MAR dropout to hold. According to Verbeke and Molenberghs (2000), these mixed-effect models permit the inclusion of subjects with dropout at some time points for both dropout patterns, namely monotone and intermittent. Since likelihood-based 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 modeling framework which can be combined with the ignorability assumption. For Y_i the vector of observation from individual i, the model can be written

as follows $Y_i = X_i\beta + Z_i b_i + \varepsilon_i$, **(22)**
where $b_i : N(0, D)$, $\varepsilon_i : N(0, \Sigma_i)$ and $b_1,...,b_N, \varepsilon_1,...,\varepsilon_N$ are independent. The meaning of each form in equation (22) are described as follows. Y_i is the n_i dimensional response vector for subject i, containing the outcomes at n_i various measurement occasions, $1 \le i \le N$, N is the number of subjects, X_i and Z_i are $(n_i \times p)$ and $(n_i \times q)$ dimensional matrices of known covariates, β is the p-dimensional vector containing the fixed effects, b_i is the q-dimensional vector containing the random effects and ε_i is a n_i dimensional vector of residual components, combining measurement error and serial correlation. Finally D is a general $(q \times q)$covariance matrix whose (i, j) th element is $d_{ij} = d_{ji}$ and Σ_i is a $(n_i \times n_i)$ covariance matrix which generally depends on i only through its dimension n_i, i.e., the set of unknown parameters in Σ_i will not depend upon i. This means marginally $Y_i : N(X_i \beta, Z_i D Z_i' + \Sigma_i)$. Thus if we define $V_i = Z_i D Z_i' + \Sigma_i$ as the general covariance matrix of Y_i, 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, Little and Rubin (1987) and Rubin (1976) stated that when MAR assumption and mild regularity conditions hold, parameters θ and ξ are independent, and that likelihood based inference is valid when the missing data mechanism is ignored. In practice, likelihood of interest is then based on the factor $f(y_i^o \mid \xi)$ (Verbeke and Molenberghs, 2000). This is referred as ignorability.

(3) Application Study

This section describes the application of the aforementioned methods for handling dropouts in longitudinal clinical trials data. The methods are applied to data from heart rate experiment, for which initially there are no actual dropouts.

(3.1) Description of the Data

Here, we have to make it clear that as our current study is essentially an application study rather than a case study, we tested the performance of the two approaches by generating dropouts from a complete data. So, our main interest was to generate a random sample of the whole dataset, and then to use it for the analysis. The data set to be analyzed in this study 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 subject 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. 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 of experiment units 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.

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, and we are also interested to investigate the differences between the drug and time effects in least-square means.

(3.2) Model Formulation

In the proceeding, we analyze the data from the clinical trial introduced above by formulating a model based on the data with heart rates. According to the study design, we include the fixed categorical effects of drug, time, and drug-by-time interaction. Therefore, the continuous outcome for the analysis reported here was heart rate, or as we will denote it in the remainder of this study, HR. Let HR_{ijk} 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 HR_{ijk}, where the response of the subject i at time j:

$$HR_{ijk} = \beta_0 + \beta_1 Time_j + \beta_2 Drug_k + \beta_3 (Time * Drug)_{jk} + \varepsilon_{ijk}, \quad \textbf{(23)}$$

where $(Time * Drug)$ denotes the drug-by-time interaction and ε_{ijk} are unknown independent and identically distributed normal random error, with mean 0 and variance σ_ε^2. As mentioned above, in this data set, there are no actual dropouts. This provides us with an opportunity to generate dropouts missing at random in order to compare the performance of MI and IPW methods to deal with dropouts.

(3.3) Generating Dropouts and the MAR Mechanism

Since there are no dropouts in the example data set described above, it provides us with an opportunity to design a comparative study to compare the two methods to deal with dropout using the results from the complete data analysis as the reference. We used the full data set to artificially generate dropouts by mimicking the MAR mechanism. From the complete data set described above, 1000 random samples of n=96 were

drawn. The dropouts in HR were created according to the MAR assumption, assuming the missingness in HR is related to observed values, in the sense that patients with higher HR at one measurement occasion tend to dropout out of the experiment at the next occasion. 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 measurement) 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 measurements. The other predictor variables other than HR, were however kept intact. For MAR, three dropout rates were implemented. The dropout rates were set at 10%, 20% and 30%. Dropouts were created in HR by randomly deleting 10%, 20%, and 30% of all measurements greater than 75 as a threshold indicating high heart rate. The observations that triggered the missing data were kept but all other subsequent observations were deleted. This scenario was generated or replicated 1,000 times. Each generated samples was analyzed using LMM and IPW to derive parameters of interest. A monotone missing pattern was assumed, which is to say that for each patient, if a HR's observation was deleted for a third time point, the subsequent observation in the fourth time point for that patient was also deleted.

(3.4) Analysis of IPW

The IPW method was applied to each generated sample using two SAS macros provided by Molenberghs and Verbeke (2005). The macros are *DROPOUT* and *DROPWGT*. These macros construct the variables "dropout" and "previous measurements" and to pass the weights (predicted probabilities) to be used for WGEE. In contrast to LMM, the IPW approach needs a model for the dropout mechanism. Therefore, for the MAR assumption, we assume the IPW models the missingness mechanism via logistic regression model, introduced in model (24), which requires the data to be MCAR or MAR. IPW was applied using to the following steps: 1) The dropout model was fitted within logistic regression using *DROPOUT* macro. The outcome variable "dropout" indicator for HR was generated, and it was binary taking the value 1 when the HR is observed, 0 otherwise, consequently indicating whether or not dropout occurred at a given time from the start of the measurement until the end of the study period (Molenberghs and Verbeke (2005)). In the dropout model, predictor variables were the outcomes at previous occasions $(y_{i,j-1})$, supplemented with genuine covariate information. To estimate the dropout probabilities, we used the following logistic regression of dropout indicators

$$logit[P(D_i = j \mid D_i \geq j)] = \psi_0 + \psi_1 y_{i,j-1} + \gamma drug_j, \quad \textbf{(24)}$$

where $y_{i,j-1}$ is the binary indicator at the previous occasion. 2) Using data and fitted probabilities from step (1), a weighted regression of the response variable in model (12) was fitted based on the inverse of the "probability of a patient dropping out at a given time and was not missing in all the previous times" as weights. This was done by using the *DROPWGT* macro in SAS. These weights were defined at the individual measurement level and were equal to the product of the probabilities of not dropping out up to the measurement occasion (Molenberghs and Verbeke (2005)). The last factor was the probability of either dropping out at that time or continuing with the study. 3) Once the selected model (24) is fitted and the weight distribution checked, we formulate the full-data regression model using inverse probability weighting. The weighted regression model is formulated by re-defining the response as $Y_{ij}^* = \hat{w}_{ij}(\alpha)Y_{ij}$ and covariate as $x_{ij}^* = \hat{w}_{ij}(\alpha)x_{ij}$. Now, let HR_{ij} denote the heart rate from patient i at time j for $j = 1,2,3,4$. Further, let x_{ij}^* be a vector of covariates with length p, where $p = 1,2,3$. Then, the mean response model can be expressed as follows

$$E(HR_{ij}^* \mid x_{ij}^*) = \mu_{ij}^* = \beta_0 + \beta_1 x_{ij1}^* + \beta_2 x_{ij2}^* + \beta_3 x_{ij3}^*, \quad \textbf{(25)}$$

where the covariate x_{ij1}^* denotes the time, the covariate x_{ij2}^* denotes the drug group, the covariate x_{ij3}^* denotes the drug by time interaction, β_0 is the population average intercept and β_1, β_2, β_3 is the average rate of change due to the time, drug main effects and their interaction. Since the response variable of interest at each occasion was the HR which is continuous, we used the identity link function and the scale parameter, ϕ. In addition to the marginal model in (25), the covariance structure of the correlated HR weights on a given patients should be modeled. In the application of IPW only first order-autoregressive AR(1) and compound symmetry covariances can be implemented. The other structures, such as unstructured covariance, toeplitz and heterogeneous (AR), may easily present computational problems. Therefore, we used the AR(1) covariance structure since it is the most reasonable in longitudinal data analysis problems. Using model (25), the parameter estimates can be calculated as the root of the weighted estimating equations

$$\sum_{i=1}^{n} = (Y_i - X_i\beta)' \Sigma_i^{-1} W_i(\hat{\alpha})(Y_i - X_i\beta) = 0, \quad \textbf{(26)}$$

where Y_i and X_i are vectors of HR and covariates, respectively, for i th patient and $W_i(\hat{\alpha})$ is a diagonal

matrix consisting of inverse probability weights for the i th patient. Model (25) was fitted using SAS procedure GENMOD with a WEIGHT statement.

(3.5) Assessment Criteria

The performance of LMM and IPW was assessed on two criteria, namely bias and efficiency. These two criteria have been used by Schafer and Graham (2002) to study the performance of list-wise deletion, single imputation, maximum likelihood and MI. Also, they used by Graham and Schafer (1999) to evaluate the the performance of multiple imputation using small multivariate data sets. Here, we defined these criteria as follows: bias refers to the differences between the average of the 1,000 coefficient estimates and the corresponding true coefficient obtained from a mixed model analysis of the original complete data. Thus, a better technique is that which does on average approach the population value with less bias. Efficiency was defined as the variability of the estimates around the true population coefficient. It was calculated by the average width of the 95% confidence interval. The 95% confidence interval width is approximately four times the magnitude of the standard error. Thus, a wider interval implies a less efficient method.

Results

The results of bias and efficiency of the drug main effect means and the time main effect means are given in Table 1 for the LMM and IPW methods.

Here we have to make it clear that we do not show full output, as the results of interaction terms are excluded. An examination of bias criterion revealed that LMM was notable for consistently producing less biased estimates vis-a-vis those estimates in IPW. It would appear that Mallinckrodt *et al,* (2003a, 2003b) recommendation to use likelihood-based analysis for incomplete longitudinal data with continuous outcomes is supported by the results presented here. In addition, this advantage of LMM is well documented in terms of continuous outcomes (see, Verbeke and Molenberghs, 2000). As a result, the bias of the estimates by IPW appeared to be independent of the dropout rates. The results based on LMM and IPW were somehow similar for 10% in some cases. They yielded estimates closer to each other. We refer here to estimates of $time_1$, $time_3$ and $time_4$. Generally the bias for LMM is negligible, regardless of the dropout rate, with some exceptions. The exceptions were with the estimates of $time_1$ and $time_4$ under 10%, as well as the estimates of $time_4$ under 20% and the estimates of BWW9 and $time_4$ under 30%. As noted above, a wider interval implies a less efficient approach, thus the widest and hence the worst, 95% confidence intervals are highlighted. Across all the three dropout rates, IPW was uniformly the worst approach in terms of efficiency, regardless of the dropout rates. Estimates which showed more efficient were $time_2$ for 20% and 30%. The LMM was more efficient most frequently. Hence, LMM was more robust than IPW against loss of efficiency due to increased dropout rate.

Table 1: Bias and Efficiency of LMM and IPW, under Different Dropout Rates: MIXED Least Squares Means (Interaction Terms are not Shown)

		Bias		**Efficiency**	
Dropout	**Effect**	**LMM**	**IPW**	**LMM**	**IPW**
	AX23	0.28	1.20	0.89	1.15
	BWW9	-0.18	-1.75	0.90	1.13
10%	CTRL	0.29	1.25	0.89	1.10
	Time1	0.50	0.48	0.96	1.61
	Time2	0.50	1.45	0.96	1.13
	Time3	0.18	0.22	1.09	1.26
	Time4	0.49	0.46	1.07	1.12
	AX23	0.41	1.40	0.93	1.14
	BWW9	0.38	1.40	0.94	1.07
20%	CTRL	0.64	1.90	0.94	1.04
	Time1	0.98	1.90	0.96	1.37
	Time2	0.48	1.33	0.99	0.78
	Time3	1.10	1.37	1.27	1.54
	Time4	0.24	0.14	1.27	1.34
	AX23	1.24	1.46	1.08	1.20
	BWW9	1.14	1.01	1.08	1.08
30%	CTRL	1.13	1.20	1.09	1.20
	Time1	0.56	1.41	0.97	1.16
	Time2	1.01	1.07	0.98	0.74
	Time3	1.05	1.27	1.55	1.68
	Time4	0.83	0.64	1.58	1.66

Note: *The largest bias and less efficiency for each given estimate presented in bold. LMM=linear mixed model; IPW=inverse probability weighting.*

Table 2: Bias and Efficiency of LMM and IPW, under Different Dropout Rates: Pairwise Comparisons among Drug Main Effect Means and Time Main Effect Means - Differences of Least Squares Means

						Bias		**Efficiency**	
Rate	**Effect**	**Drug**	**Time**	**Drug**	**Time**	**LMM**	**IPW**	**LMM**	**IPW**
	Drug	AX23		BWW9		0.47	0.55	1.27	1.61
	Drug	CTRL		CTRL		0.02	0.04	1.27	1.60
	Drug	CTRL		CTRL		-0.47	-0.50	1.27	1.58
	Time		1		2	0.28	0.97	1.37	1.89
10%	Time		1		3	0.50	0.74	1.44	2.03
	Time		1		4	0.03	0.02	1.44	1.97
	Time		2		3	0.58	0.59	1.42	1.75
	Time		2		4	0.99	0.99	1.46	1.60
	Time		3		4	0.48	0.76	1.52	1.69
	Drug	AX23		BWW9		0.24	0.73	1.33	1.56
	Drug	CTRL		CTRL		0.21	0.39	1.33	1.54
	Drug	CTRL		CTRL		0.14	0.12	1.33	1.49
	Time		1		2	0.39	0.43	1.38	1.54
20%	Time		1		3	0.68	1.02	1.57	2.02
	Time		1		4	0.15	0.24	1.59	1.93
	Time		2		3	0.54	0.76	1.67	1.84
	Time		2		4	0.89	1.10	1.62	1.54
	Time		3		4	0.63	0.94	1.83	2.09
	Drug	AX23		BWW9		0.23	0.32	1.53	1.62
	Drug	CTRL		CTRL		0.25	0.22	1.54	1.71
	Drug	CTRL		CTRL		0.26	0.37	1.53	1.63
	Time		1		2	0.49	0.82	1.41	1.36
30%	Time		1		3	0.89	1.05	1.82	2.03
	Time		1		4	0.64	0.76	1.84	2.01
	Time		2		3	0.72	0.88	1.90	1.94
	Time		2		4	0.91	1.16	1.90	1.94
	Time		3		4	1.15	1.01	2.62	2.37

Note: *The largest bias and less efficiency for each given estimate presented in bold. LMM=linear mixed model; IPW=inverse probability weighting.*

We now discuss the pairwise comparisons among drug main effect means and time main effect means by looking at Table 2, which shows the results of bias and efficiency for the differences of least squares means. Examining these results we find the following. When compared with the results based on IPW, the LMM method offered better performance across all the dropout rates. LMM contained less biased estimates as compared to IPW, except for estimates of pairwise comparisons among ($time_1$, $time_4$) under 10%, (BWW9, CTRL - $time_3$, $time_4$) under 20%, as well as (AX23, CTRL - $time_3$, $time_4$) under 30%. For 10% dropout, the estimates associated with LMM became nearly indistinguishable from those with LMM as both methods yielded similar estimates, and in one case - pairwise comparison among ($time_2$, $time_4$)-they provided the same estimates. In terms of efficiency condition investigated, table 2 has shown that as expected (see, table 2), the results were very comparable to what was found in Table 1 (see, table 1). Since a wider interval implied a less efficient, thus the widest also implies the worst, 95% is highlighted. Efficiency by LMM was better than IPW, as the later yielded larger estimates. In other words, LMM tends to have smallest estimates. Thus, LMM was more efficient than IPW most frequently, with few exceptions. Specifically, estimates which showed less efficiency under LMM were ($time_2$, $time_4$) and ($time_3$, $time_4$) for 20% and 30% rates, respectively. As a result, the degree of difference in the width of the intervals between the two methods increased with increasing dropout rate. This explains that the dropouts have a serious impact on the the performance of both methods.

Discussion and Conclusion

We have discussed the performance of using the likelihood-based and IPW methods for handling continuous outcomes, when there are dropout missing at random in longitudinal clinical trials data. Both of the methods were selected for their solid foundations on the MAR dropout mechanism, and both methods can be used for continuous measurements, however a great deal of work applying the IPW has been devoted to binary measurements data indeed. Because both methods come from two opposing schools of thought, little comparison has been done between them. Our objective was to compare them for handling incomplete longitudinal clinical trials, under three different dropout rates. From the complete data set, we generated MAR dropout. The comparison between the two methods was based on a heart rate trial data, and the estimates corresponding to likelihood-based were then compared to those obtained from IPW. The comparison was assessed using two criteria, that is; bias and efficiency.

In general, our findings favoured likelihood-based over IPW. The likelihood-based analysis using LMM consistently outperformed IPW in terms of bias and efficiency. By considering both criteria simultaneously, likelihood-based performed best under all three dropout rates when compared with the IPW approach. This is to be expected as the likelihood-based analysis does well for continuous measurements as well as for the assumption of MAR (Molenberghs and Kenward, 2007). The likelihood-based approach was less biased and considerably less variable than the IPW approach. The lower variability achieved by the likelihood-based approach makes it desirable in most statistical analyses. This agrees with the theoretical results in that IPW can be less efficient and less powerful than likelihood estimators under a well specified parametric model, see, Seaman and White (2011) and Schafer and Graham (2002). Given these results, it appears that either the likelihood-based analysis for MAR dropouts with continuous measurements are preferable to the IPW approach. The latter approach was irrespective of the type of parameter of interest, associated with greater estimation bias as well as less efficiency.

The findings further revealed that despite the mechanism of dropout was MAR, the performance of IPW was unsatisfactory. This explains that IPW has shortcomings, as shown clearly in the current analysis. This can be justified by the fact that IPW is more widely used in marginal models for discrete measurements than for continuous measurements (see, Robins *et al* 1995; Fitzmaurice *et al,* 1995). Despite these drawbacks, IPW can be considered as the longitudinal binary method of choice for the primary analysis when MAR dropout holds because of its simplicity and the ease with which it can be implemented as. Therefore, IPW might become attractive in specific circumstances. Specifically, with respect to marginal models under discrete measurements. In conclusion we submit that the use of the likelihood-based and IPW methods should be conducted with care when the longitudinal clinical trials data have dropouts in continuous measurements. Further, it appeared from the current analysis that the dropouts have a serious implication on the type of outcome. Consequently, to carefully address dropout using likelihood-based and IPW, the effect must be thoroughly explored by way of carefully designed simulation studies as well as a theoretical investigation.

References

Carpenter JR; Kenward MG; and **Vansteelandt SA** (2006) A Comparison of Multiple Imputation and Inverse Probability Weighting for Analyses with Missing Data. *Journal of the Royal Statistical Society: Series*, **169** (3): 571-584.
Available at: http://www. hbanaszak.mjr.uw.edu.pl/TempTxt/D5C9CD66d01.pdf

Fitzmaurice GM; Molenberghs G; and **Lipsitz SR** (1995) Regression Models for Longitudinal Binary Responses with Informative Drop-outs. *Journal of the Royal Statistical Society, Series B*, **57:** 691-704.

Grace Yi Y; and **Wenqing He** (2009) Median Regression Models for Longitudinal Data with Dropouts. *Biometrics*, **65** (2): 618-625.
Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1541-0420.2008.01105.x/abst

Graham JW; and **Schafer JL** (1999) On the Performance of Multiple Imputation for Multivariate Data with Small Sample Size. *In:* **Hoyle R** (ed), *Statistical Strategies for Small Sample Research,* Thousand Oaks, CA: Sage. pp 1-29.

Hogan JW; Roy J; and **Korkontzelou C** (2004) Tutorial in Biostatistics: Handling Drop-out in Longitudinal Studies. *Statistics in Medicine*, **23:** 1455-1497.

Horton NJ; and **Lipsitz SR** (2001) Multiple Imputation in Practice: Comparison of Software Packages for Regression Models with Missing Variables. *American Statistician*, **55** (3): 244-254.
Available at: http://econpapers.repec.org/article/besamstat/v_3a55_3ay_3a2001_

Horvitz DG; and **Thompson DJ** (1952) A Generalization of Sampling without Replacement from a Finite Universe. *Journal of the American Statistical Association*, **47** (260): 663-685.
Available at: https://www.researchgate.net/signup.SignUpAccountName.

Liang KY; and **Zeger SL** (1986) Longitudinal Data Analysis using Generalized Linear Models. *Biometrika*, **73** (1): 13-22.
Available at: http://biomet.oxfordjournals.org/content/73/1/13.abstract

Little RJA; and **Rubin DB** (1987) *Statistical Analysis with Missing Data*. John Wiley and Sons, Inc. New York, USA.
Available at: http://www.amazon.com/

Little RJA; and **Rubin DB** (2002) *Statistical Analysis with Missing Data*. John Wiley and Sons, New York, USA.
Available at: http://www.amazon.com/

Mallinckrodt CH; Clark WS; Carroll RJ; and **Molenberghs G** (2003a) Assessing Response Profiles from Incomplete Longitudinal Clinical Trial Data under Regulatory Considerations. *Journal of Biopharmaceutical Statistics*, **13** (2): 179-190.
Available at: http://www.tandfonline.com/doi/pdf/10.1081/BIP-120019265

Mallinckrodt CH; Sanger TM; Dube S; Debrota DJ; Molenberghs G; Carroll RJ; Zeigler Potter WM; and **Tollefson GD** (2003b) Assessing and Interpreting Treatment Effects in Longitudinal Clinical Trials with Missing Data. *Biological Psychiatry*, **53** (8): 754-760.
Available at: https://uhdspace.uhasselt.be/dspace/bitstream/1942/418/1/molg04.pdf

Milliken GA; and **Johnson DE** (2009) *Analysis of Messy Data. Design Experiments*, vol. 1. 2nd. ed. Chapman and Hall/CRC, London, UK. pp 1-655.
Available at: http://www.amazon.com/ Analysis-Messy- Data-Volume-Experiments/DP/

Molenberghs G; and **Kenward M** (2007) *Missing Data in Clinical Sudies*. John Wiley & sons, Ltd., New York, USA, pp1-506.
Available at: http://onlinelibrary.wiley.com/book/10.1002/9780470510445

Molenberghs G; and **Verbeke G** (2005) *Models for Discrete Longitudinal Data*. Springer, New York, USA, pp1-683.
Available at: http://books.google.com.bh/books/about/Models_for_Discrete_

Robins JM; Rotnitzky A; and **Zhao LP** (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.
Available at: http://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476493
Available at: http://www.questia.com/library/1G1-16679052/analysis-of-

Robins JM; and **Rotnitzky A** (1995) Semiparametric Efficiency in Multivariate Regression Models with Missing Data. *Journal of the American Statistical Association*, **90** (429): 122-129.
Available at: http://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476494
Available at: http://www.questia.com/library/1G1-16679052/ semiparametric -

Rubin DB (1976) Inference and Missing Data. *Biometrika*, **63:** 581-592.

Rubin DB (1987) *Multiple Imputation for Nonresponse in Surveys*. John Wiley, New York, USA. pp 320.
Available at: http://books.google.com.bh/books?id=bQBtw6rx_

Schafer JL (1997) *Analysis of Incomplete Multivariate Data*. Chapman and Hall, London, UK. pp 448.
Available at: http://books.google.com.bh/books?id=3TFWRjn1f- =

Schafer JL (1999) Multiple Imputation: a Primer. *Statistical Methods in Medical Research*, **8** (1): 3-15.

Schafer JL; and **Graham JW** (2002) Missing Data: Our View of the State of the Art. *Psychological Methods*, **7:** 147-177.

Schafer JL; and **Olsen MK** (1998) Multiple Imputation for Multivariate Missing-data Problems: A Data Analysts Perspective. *Multivariate Behavioral Research*, **33** (4): 545-571.
Available at: http://www.tandfonline.com/doi/pdf/10.1207/s15327906mbr3304_5

Scharfstein DO; Rotnitzky A; and **Robins JM** (1999) Adjusting for Nonignorable Dropout using Semiparametric Nonresponse Models (with discussion). *Journal of the American Statistical Association,* **94** (448): 1096-1146.
Available at: http://www.questia.com

Seaman, S R; and **White LR** (2011) Review of Inverse Probability Weighting for Dealing with Missing Data. *Statistical Methods in Medical Research*, **0:** 1-18.
Available at: http://www.sagepub. com/content/

Verbeke G; and **Molenberghs G** (2000) *Linear Mixed Models for Longitudinal Data*. Springer, New York, USA. pp569
Available at: http://books.google.com.bh/books?id=nIoME3vzND0C&printsec=

Yi GY; and **Cook RJ** (2002a) Marginal Methods for Incomplete Longitudinal Data Arising in Clusters. *Journal of the American Statistical Association*, **97** (458): 1071-1080.
Available at: http://www.questia.com/library

Yi GY; and **Cook RJ** (2002b) Second Order Estimating Equations for Clustered Longitudinal Binary Data with Missing Observations. *In:* **Chaubey YP** (ed), *Recent Advances in Statistical Methods*, World Scientific Publishing Company, London, UK. pp 352-366.
Available at: http://books.google.com.bh/books?id=WTD0bcY01ccC&dq=Recent+