Sensitivity analysis of longitudinal data With intermittent missing values: Application in a clinical trial

Abeer S. Ahmed

The National Centre for Social and Criminological Research, Cairo, Egypt

Abstract

we conduct two types of sensitivity analyses of study conclusion, in the intermittent setting: the first is fitting different distribution to the response variable. The marginal distribution of the response is assumed to be skewed distribution, the lognormal distribution in particular. The second is fitting several models for the missing data mechanism, for example, modeling the missingness based on a generalized linear model, with logit and probit link function. The model can be extended to permit possible relationships between the missing data process and covariates, for example time. The selection model for incomplete longitudinal data is presented. The stochastic EM algorithm is proposed and developed for skewed distribution model, the lognormal distribution in particular. Models for the missing data mechanism are presented. The proposed methods are applied to a data set from the International Breast Cancer Study Group.

1. Introduction

In longitudinal studies, each subject is measured repeatedly for the same response (outcome) variable either under different conditions, or different times, or both. Missing response data is a very common occurrence under such studies because of treatment drop-out, study drop-out, mistimed measurements, subjects being too sick to come to the clinic to be measured, and so forth. A subject's response can be missing at one follow-up time and then measured again at the next follow-up time, resulting in intermittent missing data patterns. Often, missing response data in these studies is nonignorable in the sense that the reason for missingness often depends on the missing values themselves. For example, the side effects of the treatment may make the patients worse and thereby affect patient participation. Suh data present

a considerable modeling challenge for the statistician. It might therefore be necessary to accommodate missingness in the modeling process.

Numerous missing data methods are formulated as selection models (Little and Rubin, 1987). A selection model factors the joint distribution of the measurement and response mechanism into the marginal measurement distribution and the response distribution, conditional on the measurements. Diggle and Kenward (1994) proposed a selection model for continuous longitudinal data with nonignorable dropout. They specified a normal linear model for the response variable and a logistic model for the probability of dropout. It is clear that the key assumptions underlying the selection model are: (i) the correct specification of the response distribution, and (ii) the missing data mechanism model. As pointed out by Diggle and Kenward themselves and many other discussants, the conclusion drawn from such a model relies on assumption which cannot be verified from the observed data. So, it is helpful to assess the impact of these assumptions on the underlying parameters through sensitivity analyses.

Several Types of sensitivity analyses have been proposed. Fitting different distributions to the response enables us to investigate the properties of the proposed method when the response distribution is misspecified. Kenward (1998) fits the normal model and the t model, in the dropout setting. Other approaches of performing sensitivity analyses are fitting several models for the missing data mechanism; see, for example, Ibrahim *et al.* (2001). Minini and Chavance (2004) suggest using a range of different values of missingnes process, which allow us to assess the sensitivity of study conclusions to dropout mechanism . Gad and Ahmed (2006) extended the Minini and Chavance's approach to the intermittent setting. Fitting the pattern-mixture model in addition to the selection model could be a sensitivity analyses tool; see, for example, Kenward and Molenberghs (1999), Thijs *et al.* (2002), Molenberghs *et al.* (2003). The local influence approach, which investigates the effect of small perturbations of missingness model on the study's conclusions is another tool; see, for example, Steen et al. (2001), Verbeke *et al.* (2001), Jansen et al. (2003), Moreno and Chavance (2016).

Most of these sensitivity analysis approaches have been applied in the dropout setting. Less attention has been paid to sensitivity analysis in the intermittent setting.

In this paper, we conduct two types of sensitivity analyses of study conclusion, in the intermittent setting: the first is fitting different distributions to the response variable. The

marginal distribution of the response is assumed to be skewed distribution, the lognormal distribution in particular. The second is fitting several models for the missing data mechanism, for example, modeling the missingness based on a generalized linear model, with logit and probit link function. Also, in many cancer clinical trials, the side effects of the treatment may affect participation, and missingness can depend on the outcome and the time covariate. In Section 2, the selection model for incomplete longitudinal data is presented. In Section 3, the stochastic EM algorithm is proposed and developed for skewed distribution model, the lognormal distribution in particular. Models for the missing data mechanism are presented in Section 4. The proposed methods are applied to a data set from the International Breast Cancer Study Group in Section 5.

2 Selection model for incomplete longitudinal data

Let $Y_i = (Y_{i1}, Y_{i2}, \ldots, Y_{im})$ be a vector containing the responses Y_{ij} for subject $i = 1, \ldots, m$ at time $j = 1, \ldots, n_i$, $X_{ij} = (X_{ij,1}, X_{iJ,2}, \ldots, X_{ij,q})$ is a covariate vector recorded for each subject at each occasion, and X_i respect the $n_i \times q$ known covariance matrix for the i^{th} individual.

Let $R_i = (R_{i1}, R_{i2}, \ldots, R_{in_i})$ be a vector described the missing data process, where R_{ij} is an indicator variable taking value 1 if Y_{ij} is observed and 0 if Y_{ij} is missing. Response vector Y_i can be divided into two vectors, based on whether values are observed or missing: $Y_i = (Y_{i,obs}, Y_{i,mis})$.

In a selection model, the joint distribution functions of Y_i and R_i is factorized as product of the marginal distribution of Y_i and the conditional distribution of R_i given Y_i. Thus

$$f(Y_i, R_i | \theta, \psi) = f(Y_i | \theta) P(R_i = r_i | Y_i, \psi),$$

where θ is a vector containing the model parameters, $P(R_i = r_i | Y_i, \psi)$ is the distribution that characterizes the mechanism, and ψ is a vector of parameters that govern the missing data mechanism.

Rubin (1976) and subsequently little and Rubin (1987) introduced the following classification of missing data mechanism. Missingness is defined to be *Missing Completely At Random* (MCAR) if Y_i and R_i are independent, i.e.

$$P(R_i = r_i | Y_{i,obs}, Y_{i,mis}, \psi) = P(R_i = r_i | \psi).$$

Missing At Random (MAR) if the conditional distribution of Ri given Yi depends only on the observed data, $Y_{i,obs}$,

$$P\left(R_i = r_i \middle| Y_{i,obs}, Y_{i,mis}, \psi\right) = P(R_i = r_i | Y_{i,obs}, \psi).$$

and *nonrandom (nonignorable)* otherwise.

In dropout pattern, Diggle and Kenward (1994) propose a selection model for longitudinal data with nonrandom dropout. They specified a normal linear model for the response variable, Y_i, and a logistic model for the probability of dropout. They suggest modelling the probability of dropout at time d_i as a function of the measurement at time d_i and the observed measurements (history); that is,

$$P\left(D_i = d_i \middle| history\right) = P_{d_i}\left(H_{d_i}, y_{d_i}; \psi\right).$$

Also they suggest using the logistic model for the dropout process as

$$\text{logit}\left\{P_{d_i}\left(H_{d_i}, y_{d_i}; \psi\right)\right\} = \psi_0 + \sum_{j=1}^{d_i} \psi_j y_{d_{i-j+1}}.$$

Troxel *et al.* (1998) adopt the Diggle and Kenward model to longitudinal data with nonrandom intermittent missing values, this approach is computationally intractable. Gad and Ahmed (2006) proposed and developed the SEM algorithm for parameter estimation in the presence of intermittent nonrandom missing values, when the responses have normal and *t* distribution.

3. The Stochastic EM algorithm for skewed distribution model, the lognormal in particular

3-1 The Stochastic EM algorithm

The SEM algorithm has been proposed by Celuex and Diebolt (1985), also in (Diebolt and Ip, 1996), as a stochastic version of the EM algorithm. The SEM algorithm overcomes the main difficulty of the EM algorithm, in some situations, by avoiding explicit calculation of the E-step. The E-step is replaced by the stochastic step (S-step) where the missing data are imputed with a single draw from the conditional distribution of the missing data given the observed data. In the M-step, the log-likelihood function of the pseudo-complete data can be maximized using standard maximization procedures. So, the algorithm involves iterating two steps, the S-step and the maximization step (M-step) for sufficient number of iterations.

The SEM algorithm can recover multimodality of the likelihood surface (Ip, 1994). The estimated parameter values corresponding to each pseudo-complete data form a Markov chain. This Markov chain converges reasonably quickly to its stationary distribution, which is unique (Diebolt and Ip, 1996). The mean of the points, ignoring the early first points as a burn-in period,

generated by the SEM algorithm can be considered as an estimate for the parameter θ. This mean is called the SEM estimate and denoted by $\tilde{\theta}$ (Diebolt and Ip, 1996).

The SEM algorithm does not provide the standard errors of the parameter estimates. Several methods have been proposed in literature to solve this problem, see, for example, Abdallah S. A., et al, (2016). The bootstrap method is used to obtain the estimated standard errors for the estimated parameters.

3-2 Inference for lognormal model

Skewed distributions are particularly common when mean values are low, variances large, and values cannot be negative, as is the case for current clinical trial. Such skewed distributions often closely fit the log-normal distribution (Aitchison and Brown 1957, Johnson et al. 1994).

In this section we apply the stochastic EM algorithm to lognormal longitudinal data. The concentration is mainly on the missing values when the pattern is intermittent and the mechanism is nonrandom. This enables us to assess the impact of the distributional assumptions on the underlying parameters. The response vector of the i^{th} subject, Y_i, is assumed to follow the log normal distribution with correlated errors. The probability density functions of the log normal distribution with parameters μ and σ is given by:

$$f(x) = \frac{1}{\sqrt{2\,\pi\sigma}\,x} \exp\left(-\frac{(\ln(x) - \mu)^2}{2\,\sigma} \right), \qquad x \in (0, \infty)$$

Assume that the missing components of Y_i are denoted as $Y_{i,mis}$ and assume that this vector is of dimension $1 \times r$ vector, i.e. $Y_{i,mis} = \left(Y_{i,mis_1}, \ldots, Y_{i,mis_r} \right)$. To implement the SEM algorithm, a sample is drawn from the conditional distribution of the missing data $Y_{i,mis} = \left(Y_{i,mis_1}, \ldots, Y_{i,mis_r} \right)$, given the observed data $\left(Y_{i,obs}, R_i \right)$. There are many ways to sample from this distribution, in this paper Gibbs sampling technique, see for example Gelfand (2000), has been used, At the $(t+1)^{th}$ iteration $Y_{i,mis}^{(t+1)} = \left(Y_{i,mis_1}^{(t+1)}, \ldots, Y_{i,mis_r}^{(t+1)} \right)$ is simulated from the full conditional distributions this iteration is executed in r sub-steps. First, $Y_{i,mis_1}^{(t+1)}$ is simulated from the conditional

distribution $f\left(Y_{i,mis_1} \middle| Y_{i,mis_2}^{(t)}, \ldots, Y_{i,mis_r}^{(t)}, Y_{i,obs}, R_i, \theta^{(t)}\right)$. Then, in the second sub-step $Y_{i,mis_2}^{(t+1)}$ is simulated from the conditional distribution $f\left(Y_{i,mis_2} \middle| Y_{i,mis_1}^{(t+1)}, \ldots, Y_{i,mis_r}^{(t)}, Y_{i,obs}, R_i, \theta^{(t)}\right)$.

In the third sub-step $Y_{i,mis_3}^{(t+1)}$ is simulated from the conditional distribution $f\left(Y_{i,mis_3} \middle| Y_{i,mis_1}^{(t+1)}, Y_{i,mis_2}^{(t+1)} \ldots, Y_{i,mis_r}^{(t)}, Y_{i,obs}, R_i, \theta^{(t)}\right)$

Now, the two steps of the SEM algorithm can be developed in the current setting as follows:

- S-Step: At the (t+1)th iteration, a sample is drawn from the conditional distribution of the missing values $Y_{i,mis} = \left(Y_{i,mis_1}, \ldots, Y_{i,mis_r}\right)$, given the observed data $\left(Y_{i,obs}, R_i\right)$, and the current parameter estimate, $\theta^{(t)}$.The full conditional distributions does not have a standard form, hence it is not possible to simulate directly from it. An accept-reject procedure is proposed for generating the missing values. The whole procedure is as follows:
  (1) Generate a candidate value, y^* from the conditional distributions $f\left(Y_{i,mis_j} \middle| Y_{i,mis_1}^{(t+1)}, \ldots, Y_{i,mis_{j-1}}^{(t+1)}, Y_{i,mis_{j+1}}^{(t)}, Y_{i,obs}, \theta^{(t)}\right)$ for j=1, 2,..., r.
  (2) Calculate the probability of missingness for the candidate value, y^*, according to the assumed dropout model, where the parameters ψ are fixed at the current value $\psi^{(t)}$. Let us denote the obtained value as P_i. The probability of missingness will be assumed to depend only on the current and the previous response values.
  (3) Simulate a random variate U from the uniform distribution on the interval [0, 1] then take $Y_{i,mis} = y^*$ if $U \le P_i$; otherwise go step 1.
- M-Step: The M-step consists of two sub-steps, the logistic step and the normal step. In the first step, the maximum-likelihood estimates of the missing data mechanism parameters are obtained. An iterative maximum-likelihood estimation approach of binary data models, see for example Collet (1991), can be used. In the normal step, the maximum-likelihood estimates of the model parameters θ are obtained using an appropriate optimization approach. The EM scoring algorithm (Jennrich and Schluchter, 1986) is used in this paper.

When implementing the SEM algorithm we need to check the convergence of the resulting chain. Several convergence diagnostics have been proposed in literature. Out of these methods, Gelman–Rubin method (Gelman and Rubin, 1992) will be used. This method is based on

generating multiple, $k \geq 2$ chains in parallel for $n = 2p$ iterations. For each chain, this method suggests starting from different points for which the starting distribution is over-dispersed compared to the target distribution. This method separately monitors the convergence of each scalar parameter of interest by evaluating the Potential Scale Reduction Factor, (PSRF), $\sqrt{\hat{R}}$ as

$$\sqrt{\hat{R}} = \sqrt{\frac{n-1}{n} + \frac{B}{nW}}$$

where B/n is the between sequence variance and W is the average of within sequence variances. This calculation depends on the last p iterations of each sequence. The convergence is achieved if the PSRF is close to one which means that the parallel Markov chains are essentially overlapping. If the PSRF is large, then proceeding further simulation may be needed.

4. Models for the Missing Data Mechanism

Several sensitivity analyses also can conduct by changing the missing data mechanisms. One possible model for the missing data mechanism is a binomial model of the form

$$P(r|Y,\psi) = \prod_{i=1}^{m}\prod_{j=1}^{n_i} \{p(r_{ij}=1|\psi)\}^{r_{ij}} \{1-p(r_{ij}=1|\psi)\}^{1-r_{ij}}, \qquad (1)$$

where $p(r_{ij}=1|\psi)$ is modelled by a logistic regression . Another missing data model use probit link function:

$$P(D_i = d_i | D_i \geq d, H_{id}, y_{id}) = \Phi(\psi_0 + \psi_1 y_{id-1} + \psi_2 y_{id}). \qquad (2)$$

Where Φ denotes the standard normal cumulative distribution function. Both models included current and the previous time point.

The model can be extended to permit possible relationships between the missing data process and covariates, for example time.

5. Application: *IBCSG data*

This paper analyzes data set concerning the quality of life among breast cancer patients in a clinical trial taken by the International Breast Cancer Study Group (IBCSG). In the IBCSG trial VI, premenopausal women with breast cancer are followed for traditional outcomes such as relapse, death and also focused on quality of life. Patients were randomized to four different chemotherapy regimens. It is intended to compare the quality of life among the four different treatments. One of the relevant determinants of quality of life was the Perceived Adjustment to Chronic Illness Scale (PACIS). This is a one-item scale comprising a global patient rating of the

amount of effort costs to cope with illness. The PACIS measured on a continuous scale from 0 to 100 where a larger score indicates that a greater amount of effort is required for the patient to cope with her illness.

Compliance was not compulsory and patients did refuse, on occasion, to complete the assessment. Even when they refused, the patients were asked to complete an assessment at their next scheduled follow-up visit. Thus, the structure of this trial results in intermittent pattern of missing data. A patient may not appear to fill the questionnaire, and it is reasonable to conjecture that the PACIS score is missing because the patient had an excellent quality of life at the time, or more likely, an extremely poor quality of life. In either case, the longitudinal outcome would be considered nonignorable.

Once enrolled on trial VI, patients were asked to complete quality of life questionnaires at baseline (pretreatment) and at three months intervals for 15 months. Hence, each questionnaire should be filled out six times. Essentially, these six time points cover the time during the administration of chemotherapy across all the four treatments. The total numbers of patients who participated in the study period is 456 patients. Ten patients who died within the study period are excluded from the analysis, so the missing responses are not due to death. The patients with missing response at the first assessment (64 cases) are excluded from the analysis; leaving 382 patients who remained alive during the 15 months of the study. The PACIS values were missing for 77% of the patients for at least one occasion, so the study completers are 89 (23%) patients. The amount of missing data increases over time, with 29%, 36%, 47%, 54% and 62% for the consecutive visits starting from the second time point. The percentages of patients with 0, 1, 2, 3, 4, and 5 missing responses were, respectively, 23%, 18%, 13%, 13%, 14% and 19%.

The analysis of this data set is challenging due to several difficulties, namely, intermittent (nonmonotone) pattern of missing data, and nonignorable missing data mechanism. Dealing with intermittent missing values is more difficult than dealing with dropouts because the missing data missing data in the intermittent pattern are sporadic and over-dispersed with respect to the response, such pattern of missing data need to be accommodated. If the data missing is nonignorable, an appropriate statistical analysis needs to take into account not only the structure of the data, but also the missing data mechanism. Methods that do not model the missing data

mechanism would be biased. Commercially available software for fitting nonignorable is not yet common.

Also, the interest in these data arises from that, several analyzes have been performed before. Hürny et al., (1992), for example, analyzed the preliminary version of these data, the responses for the first 9 months of the study, was analyzed. Only patients with complete responses are included in the analysis (complete cases analysis). Another analysis of these data has been conducted by Troxel et al. (1998), based on the responses for the first 6 months of the study. They adopt the Diggle and Kenward (1994) model to longitudinal data with nonrandom intermittent missing values. The final data, include patients who remained on the study long enough to have all assessments, were analyzed by Gad and Ahmed (2006). They generalized the Diggle–Kenward model to handle intermittent missing data pattern. Gad and Ahmed (2007) extended the same model for non-normal models, the t distribution in particular. Also, Ahmed, A. S., (2010) apply Diggle and Kenward model for estimating parameters in the normal random effects model with nonignorable intermittent missing values.

We adopt a mean model that allows each treatment to have its own effect. That is,

$$\mu_j = \beta_1 \mathrm{tr}_A + \beta_2 \mathrm{tr}_B + \beta_3 \mathrm{tr}_C \qquad \text{for } j = 1,...,6,$$

where

$$\mathrm{tr}_A, \mathrm{tr}_B, \mathrm{tr}_C = \begin{cases} (1,0,0) \text{ for treatment A} \\ (0,1,0) \text{ for treatment B} \\ (0,0,1) \text{ for treatment C} \\ (0,0,0) \text{ for treatment D.} \end{cases}$$

The first-order auto-regressive AR (1) model is adopted for the covariance structure. In this model, the $(i,j)^{th}$ element of the covariance matrix, σ_{ij} is equal to $\sigma^2 \rho^{|i-j|}$ for $j = 1,...,6$. An advantage of the proposed approach is that different covariance structures can be used.

There are main types of sensitivity analysis, the approaches considered in this paper:

- Examining sensitivity to the assumptions about the response distribution.
- Examining sensitivity to the assumptions about the missing data mechanism.

5.1 Examining sensitivity to the assumptions about the response distribution

For the missing data mechanism, we use the logistic regression model, including only the previous and the current responses to keep the model simple. That is,

$$\text{logit}\left(r_{ij}=1|\psi\right)=\psi_0+\psi_1 Y_{ij-1}+\psi_2 Y_{ij}.$$

Table 1 shows The SEM estimates and standard errors for the PACIS response of the lognormal, the normal, and the t distribution models .

Table 1
The SEM estimates and standard errors (SE) for the PACIS response

	Normal model		The t distribution model		The lognormal distribution model	
parameter	*Est.	SE	** Est.	SE	Est.	SE
tr_A	-0.20	0.17	-0,12	0.13	-0.229	0.132
tr_B	0.04	0.17	0.04	0.28	0.200	0.131
tr_C	-0.72	0.18	-0.61	0.43	-6.175	0.130
ρ	0.42	0.02	0.47	0.25	٠,532	0.03
σ^2	4.49	0.12	3.75	0.29	5.259	0.19
ψ_0	1.22	0.07	1.34	0.42	-3.598	0.104
ψ_1	1.61	0.08)	1.21	0.08	٠,235	0.032
ψ_2	1.06	0.08	0.15	0.12	٠,326	0.019
-2logL	5894		5424		5071	
AIC	5894.386		5424.386		5071.386	
BIC	5941.563		5471.563		5118.563	

We used the likelihood ratio, the corrected Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) to determine the best-fitting model. The model with the smaller likelihood ratio, AIC and BIC value is the lognormal model. So, in conclusion, assuming that the missingness model is correct, the lognormal model can be used rather than the normal model or the t model.

5.2 Examining sensitivity to the assumptions about the missing data mechanism.

Two alternative modeling strategies for the missing data mechanism are assumed data. Model 1: we use the probit regression model, including only the previous and the current responses. That is,

$$P\left(D_i = d_i \middle| \mathrm{D_i} \geq d, H_{id}, y_{id}\right) = \Phi\left(\psi_0 + \psi_1 \mathrm{Y_{ij-1}} + \psi_2 \mathrm{Y_{ij}}\right).$$

Model 2: we also construct a missing data model in which the time was included as a covariate in the missing data mechanism along with the other covariate (time). That is,

$$P\left(D_i = d_i \middle| \mathrm{D_i} \geq d, H_{id}, y_{id}\right) = \Phi\left(\psi_0 + \psi_1 \mathrm{Y_{ij-1}} + \psi_2 \mathrm{Y_{ij}} + \psi_3 \mathrm{time_{ij}}\right).$$

Table 2 shows the SEM estimates for the PACIS response when using the lognormal model assuming probit regression model under model 1 and model 2.

Table 2
The SEM estimates assuming probit regression model for the PACIS response

parameter	Model 1	Model 2
$\mathrm{tr_A}$	-0.377	-0.295
$\mathrm{tr_B}$	0.119	0.157
$\mathrm{tr_C}$	-0.672	-0.696
ρ	0.643	0.485
σ^2	6.418	4.789
ψ_0	0.044	0.0436
ψ_1	-0.021	-0.0213
ψ_2	0.127	0.126
time	-	0.2675

The results changed little if we added time into the missing data mechanism. On changing the distributional assumption from the normal to lognormal model and useing the probit regression model, the missing data mechanism moves from MNAR to MAR.

References

Ahmed, A. S. (2010), Using the Stochastic EM algorithm for the Normal Random-Effects model in longitudinal data with intermittent missing values, *The Egyptian Statistical Journal,* **54**(1).

Aitchison, J., Brown, J.A.C. (1957), The lognormal distribution, Cambridge University Press.

Celuex, G., Diebolt, J. (1985), The SEM algorithm: a probabilistic teacher algorithm derived from the EM algorithm for the mixture problems, *Computational Statistics Quarterly*, **2**, 73–82.

Collet, D. (1991), Modelling Binary Data, Chapman & Hall, London.

Diebolt, J., Ip, E.H.S. (1996), Stochastic EM: method and application, in: W.R. Gilks, S. Richardson, S., and D.J. Spiegelhalter (eds.), Markov Chain Monte Carlo in Practice. Chapman and Hall, London, Chapter 15.

Diggle, P.J., Kenward, M.G. (1994), Informative dropout in longitudinal data analysis, *Journal of Royal Statistical Society*, B **43**, 49–93.

Efron, B. (1994), Missing data, imputation, and the bootstrap. *Journal of the American Statistical Association*, **89**, 463–475.

Gad, A. M. and Ahmed, A. S., and Soliman, Abdallah. (2016), Maximum Likelihood Approach for Longitudinal Models with Nonignorable Missing Data Mechanism Using Fractional Imputation, *online in the American Journal of Applied Mathematics and Statistics*, **4(3)**.

Gad, A. M. and Ahmed, A. S.(2007), Sensitivity Analysis of longitudinal data with intermittent missing values using the Stochastic EM algorithm, *Statistical Methodology*, 4 , 217-226.

Gad, A. M. and Ahmed, A. S. (2006), Analysis of longitudinal data with intermittent missing values using the Stochastic EM algorithm, *Computational Statistics and Data Analysis*, **50**, 2702-2714.

Gelfand, A.E. (2000), Gibbs sampling, Journal of the American Statistical Association,95, 1300–1304.

Gelman, A., Rubin, D.B.(1992), Inference from iterative simulation using multiple sequences (with discussion), *Statistical* Science, 7, 457–511.

Hürny, C., Bernhard, J., Gelber, R.D., Coates, A., Gastiglione, M., Isley, M., Dreher, D., peterson, H., Goldhirsch, A., Senn, H.J.(1992), Quality of life measures for patients receiving adjutant therapy for breast cancer: an international trial, *European J. Cancer*, **28**, 118–124.

Ibrahim, J.G., Chen, M.H., Lipsitz, S.R. (2001), Missing responses in generalized linear mixed models when the missing data mechanism is nonignorable, *Biometrika*, **88**, 551–564.

Ip, E.H.S. (1994), A stochastic EM estimator in the presence of missing data: theory and applications, *Technical Report*, Division of Biostatistics, Stanford, California.

Jansen, I., Molenberghs, G., Aerts, M., Thijs, H., Steen, K.V. (2003), A local influence approach applied to binary data from a psychiatric study, *Biometrics*, **59**, 409–418.

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

Johnson NL, Kotz S, Balkrishan N. (1994), Continuous Univariate Distributions, Johne Wiley and Sons, New York.

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

Kenward, M. G., Molenberghs, G. (1999), Parametric models for incomplete continuous and categorical longitudinal data, *Statistical Methods in Medical Research*, 8, 51–83.

Little, R.J.A., Rubin, D.B. (1987), Statistical Analysis with Missing Data, Johne Wiley and Sons, New York.

Moreno and Chavance, (2016), Sensitivity analysis of incomplete longitudinal data departing from the missing at random assumption: Methodology and application in a clinical trial with drop-outs, *Statistical Methods in Medical Research*, **25(4)**, 1471-1489.

Minini, P., M. Chavance, (2004), Sensitivity analysis of longitudinal data with drop-outs, *Statistics in Medicine*, **23**, 1039–1054.

Molenberghs,G. , Thijs, H., Kenward, M.G. and Verbeke,G. (2003), Sensitivity analysis for continues incomplete longitudinal outcomes, *Statistica Neerlandica* , **57**,112–135.

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

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

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

Troxel, A.B., Harrington, D.P., Lipsitz, S.R. (1998), Analysis of longitudinal data with non-ignorable non monotone missing values, *Annals of Statistics*, **47**, 425–438.

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