Academic Journal of Applied Mathematical Sciences
ISSN(e): 2415-2188, ISSN(p): 2415-5225
Vol. 2, No. 3, pp: 19-26, 2016
URL: http://arpgweb.com/?ic=journal&journal=17&info=aims

Modeling Longitudinal Count Data with Missing Values: A Comparative Study

Fatma El Zahraa S. Salama Department of Statistics, Institute of statistical Studies and Research, Cairo University, Egypt

Ahmed M. Gad* Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt

Amany Mousa Department of Statistics, Institute of statistical Studies and Research, Cairo University, Egypt

Abstract: Longitudinal data differs from other types of data as we take more than one observation from every subject at different occasion or under different conditions. The response variable may be continuous, categorical or count. In this article the focus is on count response. The Poisson distribution is the most suitable discrete distribution for count data. Missing values are not uncommon in longitudinal data setting. Possibility of having missing data makes all traditional methods give biased and inconsistent estimates. The missing data mechanism is missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). This article compares different methods of analysis for longitudinal count data in the presence of missing values. The aim is to compare the efficiency of these methods. The relative bias and relative efficiency is used as criteria of comparison. Simulation studies are used to compare different methods. This is done under different settings such as different sample sizes and different rates of missingness. Also, the methods are applied to a real data.

Keywords: Count data; Generalized estimating equation; longitudinal data; Missing data; Missingness mechanisms; Multiple imputations; Poisson model.

1. Introduction

In longitudinal data the individuals are followed over time and thus there are multiple observations on each individual. Longitudinal data can be collected either prospectively or retrospectively. Prospective longitudinal data are collected by following the subjects through a period of time and retrospective longitudinal data are collected by getting measurements from subject according to historical records [1].

The response variable may be binary, categorical, count or continuous. For longitudinal count data the response takes only non-negative integer values, where these integers arise from counting rather than ranking. Longitudinal count data often occurs in long-term studies that concerns with the occurrence rate of a recurrent event. The set of observations for each subject may be common or not. Longitudinal count data are common in many areas such as demographic studies, epidemiologic studies and medical studies.

Longitudinal count data are different from ordinal longitudinal data that consist of integer, where the responses fall on a specific scale and only the relative ranking is important. Dealing with count data needs more attention because of many reasons:

1. The responses (errors) are not normally distributed. If we try to normalize the responses using transformation, it is difficult to deal with zeros.
2. The variance of the response variable is likely to increase with the mean.
3. Any assumed linear model might lead to the prediction of negative counts.

The aim of this article is to compare the performance of available methods to deal with count longitudinal data in the presence of missing data. The comparison is based on the relative bias and relative efficiency. The article is organized as follows. The Poisson regression model for longitudinal count data is described in Section 2. In Section 3 we briefly describe the missing data problem in longitudinal studies and describe the missingness patterns and mechanisms. Section 4 is devoted to the available methods that deal with longitudinal count data in the presence of missing values. The simulation studies are presented in Section 5 whereas the application is presented in Section 6. Finally in Section 7 we present conclusions and discussion.

*Corresponding Author

2. Poisson Regression Model for Longitudinal Count Data

The Poisson regression model is proposed by Frome, *et al.* [2]. It is the most common model to handle count data. In this model the dependent variable is non-negative integer number. This model is suitable to describe the number of events that occur over a given period of time and rare events. It has the following formula:

$$\log E(y_{ij}) = x'_{ij}\beta, \quad (1)$$

Where, y_{ij} is the response variable, x_{ij} are the covariates or design matrix and β are parameters describe the change in the log of the population average count per unit, $i=1,2, \ldots, m$ and $j=1,2, \ldots, n$. In Poisson distribution the marginal variance is equal to the marginal mean; $E(y) = \mathrm{Var}(y)$. This model belongs to a wider class of models; the generalized linear model. The generalized linear model popularized by McCullagh and Nelder [3] which consists of three components; the random component, the systematic component and the link function. **The random component** represents the response variable and its probability distribution. In our case the probability distribution is Poisson. **The systematic component** represents the predictors (X variables) in the model. These predictors might be continuous and/or categorical and interactions between predictors. Also, polynomial function of predictors can be used. **The link function** links the random and the systematic components, and links the expected value of *Y* to the predictors. For Poisson model the log function is the link function.

3. Missing Data in Longitudinal Studies

The presence of missing values complicates the analysis and affects the properties of the estimates. There are two different patterns of missing data; dropout and intermittent. Intermittent missingness occurs whenever a subject is observed even after a missing value, while dropout is defined if the existence of the missing value indicates that the subject withdraws from the study. When the missing pattern is intermittent it is easy to find the reasons for the missingness as the subject still in the study. A dropout occurs if an individual skips a particular visit and then never comes back for subsequent visits.

The missing data mechanisms are missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). If we define an indicator variable R_{ij} to be 1 if the observation y_{ij} is available and equal 0 if y_{ij} is missing. Following Rubin's taxonomy [4] data are said to be missing completely at random when the probability of missingness is independent of both observed and unobserved data; $P(R_i|Y_i^o, Y_i^m) = P(R_i)$, where $R_i = (R_{i1}, R_{i2}, R_{i3}, \ldots, R_{in})$. The Y_i^o denote the vector of observed responses and Y_i^m denote the vector of missing responses for subject i. Data are said to be missing at random when the probability of missingness depends on the set of observed response (Y_i^o); $P(R_i|Y_i^o, Y_i^m) = P(R_i|Y_i^o)$. Data are said to be missing not at random when the probability of missingness is related to the values that should have been obtained, in addition to the one actually obtained. This relation cannot be ignored when we make inference, so this type of missing data is called non-ignorable missingness. When the data has non-ignorable missingness all standard methods of longitudinal data may not be valid.

4. Methods for Longitudinal Count Data with Ignorable Missingness

In the presence of missingness inference using traditional methods may result in invalid results. So, methods that take into account the missingness mechanism are needed. The following are the common methods that used in the case of ignorable missingness.

4.1. Complete Case Analysis (CC)

It is the most common method of dealing with missingness in the covariates or the response. In this method the observations of the participants who have any missing data are omitted. This method is simple and can be implemented using common software. However, the method has the following disadvantages:

1. There is nearly always a substantial loss of information.
2. When the missing data are not MCAR the result from the CC may be biased.
3. Dependence on the complete case can be unrepresentative of the full population and this gives an impact on reduced statistical precision and power.

4.2. Generalized Estimating Equations Method (GEE)

Liang and Zeger [5] proposed this method to study the population-average effect. The GEE approach is an extension of generalized linear models. The method is considered as a semi-parametric approach and can be used for categorical, count, and continuous response.

Let $S_i(\alpha)$ be $j \times j$ symmetric working correlation matrix completely described by the parameters vector α of length *m*. Let $V_i = \emptyset A_i^{\frac{1}{2}} S_i(\alpha) A_i^{\frac{1}{2}}$ be the corresponding working covariance matrix of Y_i, where A_i is a diagonal matrix with entries a_{ij}. For given estimates $(\hat{\emptyset}, \hat{\alpha})$ of $(\emptyset,\alpha)$ the estimate $\hat{\beta}$ is the solution of the following equations:

$$\sum_{i=1}^{N} \frac{\partial \mu_i'}{\partial \beta} V_i^{-1}(Y_i - \mu_i) = 0.$$

This scheme yields consistent estimate of β. Moreover, $N^{1/2}(\hat{\beta} - \beta)$ is asymptotically multivariate normally distributed with zero mean and covariance matrix $\Sigma = \lim_{N\to\infty} N \sum_0^{-1} \sum_1 \sum_0^{-1}$, where

$\sum_0=\sum_{i=1}^{N}\frac{\partial\mu_i'}{\partial\beta}V_i^{-1}\frac{\partial\mu_i}{\partial\beta}$, $\sum_1=\sum_{i=1}^{N}\frac{\partial\mu_i'}{\partial\beta}V_i^{-1}\mathrm{Cov}(Y_i)V_i^{-1}\frac{\partial\mu_i}{\partial\beta}$.

Replacing the parameters β, ∅ and α by consistent estimates and the covariance matrix $\mathrm{Cov}(Y_i)$ by $(Y_i-\mu_i)(Y_i-\mu_i)'$, this is called sandwich estimate $\hat{\Sigma}$ of Σ. The estimate $\hat{\Sigma}$ is a consistent estimate of Σ even if the working correlation matrices $S_i(\alpha)$ are misspecified. However, for a small number of clusters ($N \leq 30$) the sandwich variance estimator exhibits bias. Paik [6] recommended using the jackknife variance estimator which is defined as

$$\frac{N-p}{N}\sum_{i=1}^{N}(\hat{\beta}_{-i}-\hat{\beta})(\hat{\beta}_{-i}-\hat{\beta})',$$

Where p is the number of parameters in the mean structure and $\hat{\beta}_{-i}$ are the estimates of β leaving out the i^{th} cluster.

4.3. Weighted Generalized Estimating Equations (WGEE)

The weighted GEE procedure implements the inverse probability-weighted method to account for dropouts under the MAR assumption [7]. If the missing values are intermittent for any of the subjects, then the weighted generalized estimating equation method does not apply. The estimation of this method is similar to the estimation that obtained from the complete-case analysis as a solution to the quasi-likelihood estimating equations [8]. The weighted GEE procedure implements two different weighted methods (observation-specific and subject-specific) for estimating the regression parameter β occur MAR mechanism and dropout pattern. Both of the weighted methods provide consistent estimates if the missingness is MAR.

a. Observation-Specific Weighted GEE Method

Suppose w_{ij} is the weight for y_{ij}, which is defined as the inverse probability of observing y_{ij} . In other words, $w_{ij}^{-1}=P(R_{ij}=1|X_i,Y_i)$. Suppose that W_i is a j×j diagonal matrix whose j^{th} diagonal is $R_{ij}w_{ij}$. Note that R_{ij} is an indicator variable with value 1 if y_{ij} is observed and value 0 if y_{ij} is missing. The weighted generalized estimating equations [7] are given as

$$\sum_{i=1}^{N}\frac{\partial\mu_i'}{\partial\beta}V_i^{-1}W_i(Y_i-\mu_i(\beta))=0\ .$$

The weighted generalized estimating equations estimates are unbiased assuming MAR. If the observations are appropriately weighted it leads to consistent estimates of β. The weights w_{ij} are often unknown in practice and are estimated by a logistic regression model under the MAR assumption. Let $h_{ij}=P(R_{ij}=1|r_{ij-1}=1,Y_i,X_i)$ denotes the probability of observing the response y_{ij} given its observed previous responses. Under the MAR assumption,

$$h_{ij}=P(R_{ij}=1|R_{ij-1}=1,Y_i,X_i)=P(R_{ij}=1|R_{ij-1}=1,X_i,Y_1,\dots,Y_{j-1}).$$

Using the observed data, h_{ij} can be predicted from a logistic regression model,

$\mathrm{logit}\{h_{ij}\}=z_{ij}\alpha$,

Where z_{ij} are predictors that usually include the covariates x_{ij}, the past responses, and the indicators for visit times. The dropout process implies that the estimated probability of observing y_{ij}can be expressed as a cumulative product of conditional probabilities:

$$P(R_{ij}=1|Y_i,X_i)=h_{i1}(\hat{\alpha})\times\ h_{i2}(\hat{\alpha})\times\dots\times\ h_{ij}(\hat{\alpha})$$

With the estimated weights $\hat{w}_{ij}^{-1}=\hat{P}(R_{ij}=1|Y_i,X_i)$ the regression parameter β is estimated by solving the equation for S (β) after plugging in the estimated weights.

b. Subject-Specific Weighted GEE Method

This method assigns a single weight to each subject. In other words, all the observations from the same subject receive the same weight. The subject-specific weighted method obtains the regression parameter estimates by solving the equations:

$$\sum_{i=1}^{N}\frac{\partial\mu_i'}{\partial\beta}V_i^{-1}w_i[Y_i-\mu_i(\beta)]=0,$$

Where the responses for the i^{th} subject are $Y_i=(y_{i1},y_{i2},\dots,y_{ij})'$ and the weights w_i is the inverse probability that the subject i dropout at the observed time. The estimates are unbiased when the subjects are appropriately weighted and lead to consistent estimates of the regression parameters β [9].

The subject-specific weights can be estimated as a cumulative product of conditional probabilities if the dropout time (m_i) is less than or equal T (the time of last observation) as:

$$\hat{w}_i^{-1}=\Pr(r_{in_i}=0,r_{in_i-1}=1|X_i,Y_i)=[\lambda_{i1}(\hat{\alpha})\times\dots\times\lambda_{in_{i-1}}(\hat{\alpha})\times(1-\lambda_{in_i}(\hat{\alpha}))]^{-1}.$$

If $m_i=T+1$

$$\hat{w}_i^{-1}=\Pr(r_{in_i-1}=1|X_i,Y_i)=[\lambda_{i1}(\hat{\alpha})\times\lambda_{i2}(\hat{\alpha})\times\dots\times\lambda_{in_{i-1}}(\hat{\alpha}))]^{-1}.$$

Thus, the subject-specific weights $\hat{w}_i$ can be obtained depending on λ_{ij} which can beestimated by fitting a logistic regression to the data (r_{ij}, z_{ij}). The regression parameter β from the subject-specific weighted GEE method can be estimated by solving for S(β) after plugging in the estimated weights.

c. Steps to Find Parameters for the WGEE

The steps of the observation-specific or the subject-specific weighted GEE method can be summarized as follows:

1. Fit a logistic regression to the data r_{ij}, z_{ij} to obtain an estimate of α and estimate the weights.
2. Compute an initial estimate of α using ordinary generalized linear model and assuming independence of the responses.
3. Compute the working correlation matrix *S* based on the standardized residuals, the current estimate of β, and the specified structure of *S*.
4. Compute the estimated covariance matrix

$V_i=\boldsymbol{\phi}\, A_i^{\frac{1}{2}}\, \hat{S}(\alpha)\, A_i^{\frac{1}{2}}$.

5. Update $\hat{\beta}$ as

$$\hat{\beta}_{r+1}=\hat{\beta}_r+\left[\sum_{i=1}^{N}\frac{\partial\mu_i'}{\partial\beta}\,V_i^{-1}\,\frac{\partial\mu_i}{\partial\beta}\right]^{-1}\left[\sum_{i=1}^{N}\frac{\partial\mu_i'}{\partial\beta}\,V_i^{-1}\,W_i(Y_i-\mu_i)\right].$$

4.4. Imputation Methods

The imputation methods have two categories; single imputation and multiple imputations. The multiple imputations method is common in the count data setting. The multiple imputations by chained equations (MICE) is a special case of multiple imputation techniques [10]. The MICE is sometimes called "fully conditional specification" or "sequential regression multiple imputation". The MICE is very flexible and can be applied for different responses (continuous, categorical, binary and count).

In the MICE procedure a series of regression models are fitted where each variable with missing data is modeled conditional upon the other variables in the data. The chained equation process can be broken down into general steps:

1. A simple imputation is performed for every missing value in the dataset, say mean imputation.
2. The observed values from the variable in step 1 are regressed on the other variables in the imputation model, which may or may not consist of all of the variables in the dataset.
3. The missing values for the variable are then replaced with predictions (imputations) from the regression model.
4. Then, this variable is subsequently used as an independent variable in the regression models for other variables, both the observed and the imputed values are used in this case.

Steps 1– 4 are conducted for every variable has missing data. Steps are repeated for a number of iterations where imputations are updated at each iteration. At the end of these iterations the final imputations are retained, resulting in one imputed dataset. Generally, ten iterations are performed [11]. Different MICE software packages vary somewhat in their exact implementation of this algorithm (e.g., in the order in which the variables are imputed), but the general strategy is the same.

5. Simulation Study

The aim of this simulation is to compare the performance of the three methods; the GEE, the WGEE and the MICE. Data were simulated according to the following model:

$$E(y_{ij}) = \exp(\beta_0 + \beta_1 * t_{ij} + \beta_2 * trt_{ij} + \beta_3 * t_{ij} * trt_{ij}),$$

where t_{ij} represent variable (time) that takes the values 0, 1, 2, 3, 4; trt_{ij} represent dummy-variable placebo (0) and treatment (1). The number of subjects in each group is the same. The sample size is fixed at 20 (small sample size), 50 (moderate sample size), and 100 (large sample size). The missing data are generated as dropout under two types of missingness; MCAR and MAR. Two missingness rate where considered; low missing rate with missing percentage ranges from 10% to 20%, and high missing rate with missing percentage ranges from 50% to 70% approximately. The true parameter values were fixed at β_0=1.5, β_1= - 0.5, β_2= - 0.5, β_3= 0.5. The three methods were applied and the weighted GEE is based on the observation weight because it is more powerful than subject weights.

Table-1. The relative bias (RB%) of different methods at different sample sizes.

			MCAR Mechanism			**MAR Mechanism**		
n	Parameter	Missing Rate	GEE	WGEE	MI-GEE	GEE	WGE	MI-GEE
20	β_0	low	-0.5	2.6	-2.8	8.4	0.4	-3.4
	β_1		2.9	2.9	-4.0	26.4	-1.2	-7.2
	β_2		1.2	-16.0	10.7	24.8	-15.4	-16.8
	β_3		2.6	0.3	-26.1	17.2	0.4	-27.0

20	β_0	high	-0.2	15.2	-1.8	-7.4	-3.7	-3.0
	β_1		0.8	17.1	-9.2	-16.6	2.2	-2.8
	β_2		0.6	-25.5	-12.4	-19.4	-11.0	-7.0
	β_3		0.6	-2.8	-23.4	-16.8	- 9.6	-9.8
50	β_0	low	-0.2	-5.1	-3.2	-6.5	0.5	-2.9
	β_1		0.7	5.2	-3.3	-5.2	-0.2	-3.8
	β_2		0.6	-13.1	-15.0	-13.6	-2.8	-16.6
	β_3		0.7	4.8	-25.9	-1.4	0.2	-25.4
	β_0	high	-0.2	7.0	-1.8	-5.5	-0.9	-0.7
	β_1		0.8	-12.3	-12.7	-12.8	-4.4	-2.6
	β_2		0.6	23.7	-12.5	-14.8	6.6	-7.6
	β_3		0.6	0.7	-23.3	-13.2	-8.2	-10.4
100	β_0	low	-0.1	-13.8	-0.5	6.0	0.1	-2.9
	β_1		0.3	-2.4	-4.9	8.4	-0.2	-0.8
	β_2		0.5	-16.4	-4.9	-9.6	-0.2	-12.8
	β_3		0.3	-9.6	-8.7	9.0	0.0	-18.2
	β_0	high	-0.1	4.9	-1.3	-3.5	-0.1	-0.5
	β_1		0.4	-3.9	-8.5	-14.8	-1.4	-1.8
	β_2		0.04	28.1	-10.9	-10.0	-4.4	-5.8
	β_3		0.3	7.0	-16.4	-8.2	-6.2	-8.8

The comparisons are made according to the relative bias (RB%) and the relative efficiency (RE%). The RB can be obtained using the formula;

RB=$\frac{Estimate-True\ value}{true\ value} \times 100$.

The relative efficiency (RE%) can be obtained as

$$RE = \frac{var(\beta) for\ new\ method}{var(\beta) for\ complete\ method} \times 100$$

The results are shown in Table 1. We present the relative bias only for the sake of space limitation. From the results, in the missing completely at random (MCAR), we can see that all methods produce relatively unbiased estimates. The relative biased is below 30% for all methods. There is small change in the result when the missing rates and sample sizes change. Under MCAR the GEE less biased. Under MAR the WGEE is better in most of the cases than other methods; the generalized estimating equation and the multiple imputations. The WGEE has less relative bias especially when the simple size is large =100.

5.1. Application (Epileptic's Data)

This data set consists of 59 epileptics patients. The data have been analyzed by Thall and Vail [12] and Breslow and Clayton [13] as a complete data. In this study 31 patients receive the anti-epileptic drug (progabide) and the other 28 patients receive (placebo). For each patient, the number of epileptic seizure was recorded during a base line period of eight weeks. Patients were then randomized to treatment with the anti-epileptic drug group or to placebo. The number of seizures was recorded during four consecutive two-week intervals.

Hence the response Yij is the number of epileptics of subject i during the period j. The treatment groups are coded as 1 for treatment and 0 for placebo. The time points are labeled as 1, 2, 3, and 4. A base line variable is the number of seizure for every patient before the treatment. The age (in years) of patients during this study is considered.

Figure-1. Boxplot of seizures for completers for placebo and treatment

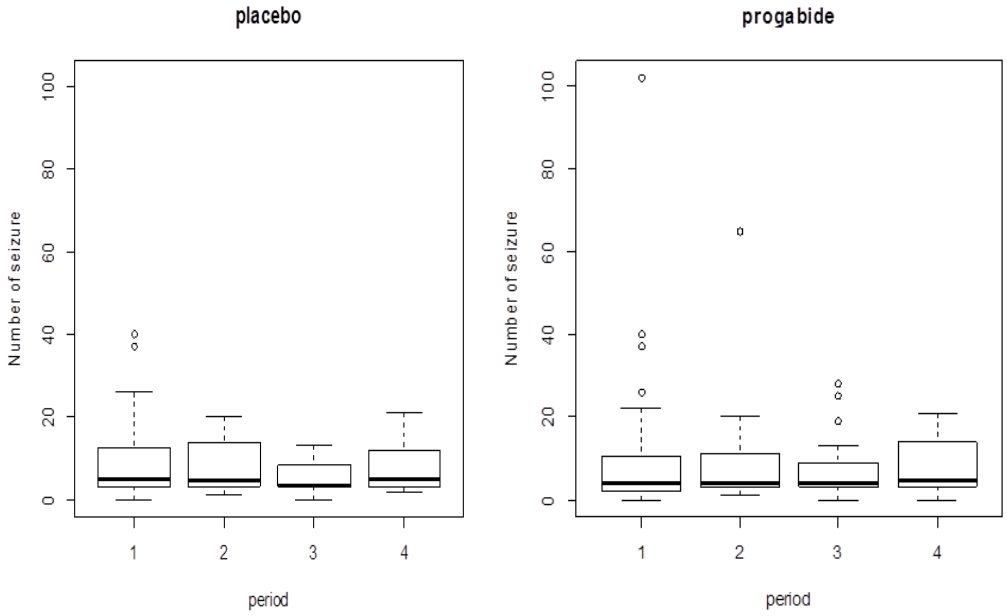


Figure 1 presents a boxplot of the number of seizures for completers in the two groups. From the Figure it can be seen that there are some outliers in the treatment group more than the placebo group. Also it can be seen that the change in the placebo group are stable over the course of the trail while for the Progabide group there is a decline in the number of observed seizure.

Table-2. Estimates and their standard errors of epileptic's data

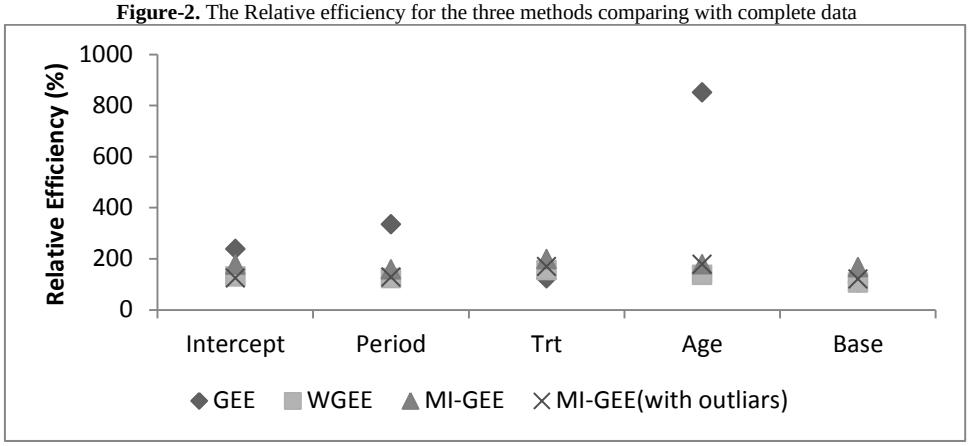
Models	Intercept	Period	Trt	Age	Base
Complete data					
Estimate	0.676*	-.059*	-0.148	0.024*	0.023*
Stander Error	0.354	0.35	0.169	0.012	0.031
Generalized Estimating Equation					
Estimate	0.267	-0.028	-0.219	0.033*	0.024*
Stander Error	0.547	0. 65	0.187	0.035	0.032
Relative Efficiency	238.76%	344.89%	122.43%	850.69%	106.55%
Weighted Generalized Estimating Equation					
Estimate	0.675	-0.051	-0.14	0.021*	0.023*
Stander Error	0.404	0.39	0.211	0.014	0.032
Relative Efficiency	130.24%	124.16%	155.88%	136.11%	106.55%
Multiple imputation- Generalized Estimating Equation					
Estimate	0.548	-0.013	-0.244	0.03*	0.021*
Stander Error	0.47	0.44	0.238	0.016	0.04
Relative Efficiency	176.27%	158.04%	198.32%	177.77%	166.49%
Multiple imputation- Generalized Estimating Equation(no outliers)					
Estimate	0.612	-0.043	-0.153	0.035*	0.048*
Stander Error	0.395	0.396	0.22	0.016	0.034
Relative Efficiency	124.5%	128.01%	169.46%	177.77%	120.29%

*: result is significant at 0.05 level

The parameter estimates and their standard errors are presented in Table 2. From the results it is clear that the values of the WGEE and multiple imputation are more close to each other than the GEE. Also for these two methods, only the age and the base variables are significant and have positive relation with the mean number of seizures.

From Figure 2 it is noted that the relative efficiency for the estimated parameters from WGEE near to 100% so this mean that it is more efficient than the other methods and GEE is the least efficient. When we remove the outliers from the data the performance of the MI-GEE became better and the WGEE and MI-GEE approximately gave the same results.

Figure-2. The Relative efficiency for the three methods comparing with complete data



6. Conclusion

Several routes are available to analyse incomplete non-Gaussian (e.g., binary, count) longitudinal data. Because the standard generalized estimating equations are unbiased except under MCAR, variety of modifications and alternatives to GEE have been proposed. An important route is through weighted estimating equations, as proposed by Rotnitzky and Robins [7]. A combination of GEE and multiple imputation methods (MI-GEE) provides an alternative route.

Although likelihood methods are appealing because of their flexible properties, their use for non-Gaussian outcomes can be problematic due to prohibitive computational requirements. Therefore, the GEE is a good alternative method especially in the case of discrete variables. Generalized estimating equations are useful to circumvent the computational complexity of full likelihood, can be considered whenever interest is restricted to the mean parameters (treatment difference, time evolutions, effect of baseline covariates, etc.). It is rooted in the quasi-likelihood ideas expressed by Nelder and McCullagh [8].

The GEE yields biased estimates when the missing data is related to the dependent variable (MAR). This study depends on another method that adds weights; the observation weights or subject weights. The WGEE needs correct specification of the dropout model. While imputation-based methodology needs a correctly specified imputation model. While each of the two methods, the WGEE and the MI-GEE, depends on the GEE analysis a comparison between approaches, inverse probability weighting and MI-GEE are considered. The behavior of both methods in terms of efficiency and bias are studied.

The WGEE is better in most of MAR cases than generalized estimating equation and multiple imputations. It has the least relative biased and more efficiency, especially when the simple size is large (n=100). Also the performance of the WGEE method is better when the rate of missingness is low. Multiple imputations have results similar to the WGEE in the case of low missingness rate. But if there are any outliers in the data, the researcher should be conservative to multiple imputation (MICE); although it is attractive and used in different patterns of missing data and any types of distribution, but it depends on the mean to find the imputed value and this is affected by outliers so the result will not be good in this case especially in the case of high missing rates.

References

[1] Goldfar, N., 1960. "An optimum property of regular maximum likelihood estimation." *Annals of Mathematical Statistics*, vol. 31, pp. 1208-1212.

[2] Frome, E. L., Kutner, M. H., and Beauchamp, J. J., 1973. "Regression analysis of poisson-distributed data." *Journal of American Statistical Association*, vol. 68, pp. 935-940.

[3] McCullagh, J. A. and Nelder, J., 1982. *Generalized linear models*. London: Chapman and Hall.

[4] Rubin, D. B., 1976. "Inference and missing data." *Biometrics*, vol. 63, pp. 581-592.

[5] Liang, K. Y. and Zeger, S. L., 1986. "Longitudinal data analysis using generalized linear models." *Biometria*, vol. 73, pp. 13-22.

[6] Paik, M. C., 1988. "Repeated measurement analysis for nonnormal data in small samples." *Communications in Statistics - Simulation and Computation*, vol. 17, pp. 1155–1171.

[7] Rotnitzky, A. and Robins, J. M., 1995. "Semiparametric efficiency in multivariate regression models with missing data." *Journal of American Statistical Association*, vol. 90, pp. 122-129.

[8] Nelder, J. and McCullagh, P., 1989. *Mathematical statistics of generalized linear model*. London: Chapman and Hall.

[9] Rathouz, P. J., Satten, G. A., and Carroll, R. J., 2002. "Semiparametric Inference in Matched Case-Control Studies with Missing Covariate Data." *Biometrika*, vol. 89, pp. 905-916.

[10] Buuren, S. V., Boshuizen, H. C., and Knook, D. L., 1999. "Multiple imputation of missing blood pressure covariates in survival analysis." *Statistic in Medicine*, vol. 18, pp. 681-694.

[11] Raghunathan, A., Lakshminarayana, G., Potlapally, N., and Ravi, S., 2002. "Algorithm exploration for efficient public-key security processing on wireless handsets." *Le Palais des Congres*, pp. 42-46.

[12] Thall, P. F. and Vail, S. C., 1990. "Some covariance models for longitudinal count data with overdispersion." *Biometrics*, vol. 46, pp. 657-671.

[13] Breslow, N. E. and Clayton, D. G., 1993. "Approximate inference in generalized linear mixed models." *Journals of the American Statistical Association*, vol. 88, pp. 9-25.