

Simulation-based study comparing multiple imputation methods for
non-monotone missing ordinal data in longitudinal settings

Peer-reviewed author version

Donneau, A.F.; Mauer, M.; Lambert, Philippe; MOLENBERGHS, Geert & Albert, A.
(2014) Simulation-based study comparing multiple imputation methods for
non-monotone missing ordinal data in longitudinal settings. In: Journal of
Biopharmaceutical Statistics, 25 (3), p. 570-601.

DOI: 10.1080/10543406.2014.920864

Handle: <http://hdl.handle.net/1942/17789>



Simulation-based study comparing multiple imputation methods for non-monotone missing ordinal data in longitudinal settings [Link](#)

Peer-reviewed author version

Made available by Hasselt University Library in [Document Server@UHasselt](#)

Reference (Published version):

Donneau, A.F.; Mauer, M.; Lambert, Philippe; Molenberghs, Geert & Albert, A.(2014)
Simulation-based study comparing multiple imputation methods for non-monotone missing ordinal data in longitudinal settings. In: Journal of Biopharmaceutical Statistics, 25 (3), p. 570-601

DOI: 10.1080/10543406.2014.920864

Handle: <http://hdl.handle.net/1942/17789>

Simulation-based study comparing multiple imputation methods for non-monotone missing ordinal data in longitudinal settings

A.F. Donneau Medical Informatics and Biostatistics, School of Public Health,
University of Liège, Liège, Belgium

M. Mauer EORTC Headquarters, Departments of
statistics and quality of life, Brussels, Belgium

Ph. Lambert Institute of Social Sciences, Quantitative Methods
University of Liège, Liège, Belgium

G. Molenberghs I-BioStat, University of Hasselt, Diepenbeek, Belgium
I-BioStat, Katholieke University of Leuven, Leuven, Belgium

A. Albert Medical Informatics and Biostatistics, School of Public Health,
University of Liège, Liège, Belgium

Version 14-12-2012

Abstract

The application of multiple imputation (MI) techniques as a preliminary step to handle missing values in data analysis is well established. The MI methods can be classified into two broad classes, the joint modeling and the fully conditional specification approaches. Their relative performance for longitudinal ordinal data setting is not well documented. This paper intends to fill this gap by conducting a large simulation study on the estimation of the parameters of a longitudinal proportional odds model. The two MI methods are also illustrated on a real dataset of quality of life in a cancer clinical trial.

Keywords: ordinal variables; longitudinal analysis; non-monotone; intermittent; missing at random; multiple imputation

Address for correspondence: A.F. Donneau, Medical Informatics and Biostatistics, School of Public Health, University of Liège, Sart Tilman B23, 4000 Liège, Belgium. E-mail: afdonneau@ulg.ac.be

1 Introduction

In clinical trials, it is common practice to assess quality of life (QoL) on a Likert-type scale along with the patient's disease evolution [1]. Patients however may withdraw prematurely from the trial or miss one or more follow-up visits. The latter situation refers to intermittent or non-monotone missingness pattern and the former to monotone missingness. The statistical analysis of non-Gaussian longitudinal data with non-monotone missingness pattern is difficult to handle. Even when the number of patients with intermittent missing data is small, discarding these patients from the analysis [2] is unsatisfactory and alternative methods have to be considered. Multiple imputation (MI) has become a reference method for handling missing data [3]. For longitudinal ordinal data with monotone missingness patterns, MI consists in a sequential application of the proportional odds model considering the previous assessment time as covariate and accounting for the uncertainty about the regression coefficients [4]. We shall refer to this method as the ordinal imputation model (OIM). Even if inappropriate for ordinal data, it is common practice to impute ordinal data using a MI approach for continuous data based on multivariate normality [5]. This MI method will be referred to as multivariate Normal imputation (MNI). In a previous work of our group, we compared the performance of both approaches for the monotone setting and we clearly demonstrates the superiority of the OIM approach [6]. The OIM method however hardly works for non-monotone missing data and it has been suggested to apply the MNI method based on multivariate normality [5] even if inappropriate for ordinal data. Here, we propose to adapt the OIM method to longitudinal ordinal data with non-monotone missingness patterns.

Multivariate MI methods can be classified into two broad classes, respectively the joint

modeling (JM) and the fully conditional specification (FCS). The latter is also known as chained equation, variable-by-variable imputation or regression switching. Within the JM approach, the joint distribution of the data has to be specified (e.g. normality). The idea of the FCS imputation method is to bypass the definition of the joint distribution by specifying a conditional distribution for each variable where data need to be imputed. In the subsequent, we shall assume that covariates are fully observed and only the ordinal outcome can be missing. Thus, a proportional odds model needs to be specified at each assessment time point.

We shall adapt the FCS strategy to monotone and non-monotone missing ordinal data by means of widely available statistical software procedures. The performance of the proposed method was compared to the joint modeling that assume a multivariate normal distribution method by focusing on the estimation of the parameters of a longitudinal proportional odds model. Both imputation methods were assessed through Monte Carlo simulated artificial data sets and also illustrated on a real example. The simulations will cover well-balanced outcome data but also skewed distributions, as often observed in QoL studies.

The paper is organized as follows. The proportional odds model to analyze longitudinal ordinal data is briefly reviewed in Section 2, while a general overview of the problem of missing data is given in Section 3. Section 4 outlines the theoretical background of multiple imputation including those for continuous and ordinal variables. The simulation experimental design is described in Section 5 and results are presented in Section 6. Both MI methods are illustrated on a QoL dataset in Section 7. Concluding remarks are given in Section 8.

2 The QoL dataset

The QoL data used in this work were obtained from the EORTC phase III clinical trial 26981 comparing radiotherapy (RT) and radiotherapy plus concomitant daily temozolomide, followed by adjuvant temozolomide (RT+TMZ) in patients with newly diagnosed and histologically confirmed glioblastoma. Between August 2000 and March 2002, a total of 573 patients were randomized by 85 institutions in 15 countries in this trial, respectively 286 in the RT arm and 287 in the RT+TMZ arm. Clinical and QoL results have been published previously [7, 8].

Per protocol, QoL had to be assessed in all patients using the EORTC QLQ-C30 version 2 questionnaire [9]. In the RT arm, QoL assessment was performed at baseline (ie, before start of treatment), during radiotherapy at 4 weeks, 4 weeks after completion of the radiotherapy and then every three months until disease progression. In the RT+TMZ arm, QoL assessment was performed at baseline, during radiotherapy and concomitant chemotherapy at week 4, 4 weeks after RT at the end of the third and sixth cycle of adjuvant temozolomide, and then every 3 months until disease progression. At the time of the analysis, time windows for acceptable QoL forms were defined around each time point to gather the maximum information available [8]. Since there were only a few assessments available after the first two follow-up time points, the analysis was stopped there.

In this paper, we shall consider the appetite loss (AP) scale of the QLQ-C30 as the outcome variable. AP is an ordinal variable with 4 response categories ('Not at all', 'A little', 'Quite a bit', 'Very much'). Since only few patients reported category 'Very much', the two last categories were combined into a single one. In the following, the time of AP assessment was treated as a categorical covariate. The distributions of AP according to time points and treatment groups are displayed in Table 1.

Table 1: Distribution of appetite loss (Number (%)) for each time point and treatment arm

Time	RT arm			RT+TMZ arm		
	Not at all	A little	Quite a bit Very much	Not at all	A little	Quite a bit Very much
T0 - Baseline	201 (81.4)	35 (14.2)	11 (4.45)	206 (85.5)	21 (8.71)	14 (5.81)
T1 - During RT	148 (78.7)	28 (14.9)	12 (6.38)	133 (66.2)	41 (20.4)	27 (13.4)
T2 - After RT	104 (73.2)	27 (19.0)	11 (7.75)	109 (66.1)	39 (23.6)	17 (10.3)
T3 - FU1	45 (73.8)	13 (21.3)	3 (4.92)	58 (62.4)	22 (23.7)	13 (14.0)
T4 - FU2	25 (80.7)	4 (12.9)	2 (6.45)	61 (75.3)	17 (21.0)	3 (3.70)

FU1 = first follow-up / FU2 = second follow-up

In cancer trials, the drop-out is typically linked to disease progression and death.

Furthermore, it has been shown that no sharp increase or decrease was observed in scores just before missingness, which is usually a good indicator for non-ignorable missing data [7, 8]. A total of 29 different missingness patterns was observed for AP. The distribution of the complete, monotone and non-monotone missingness patterns in each treatment group is summarized in Table 2.

Table 2: Distribution of the different missingness patterns (Number (%)) in both treatment arms

Missingness pattern	RT arm	RT+TMZ arm
Complete	15 (5.62)	30 (11.2)
Monotone	200 (74.9)	138 (51.3)
Non-monotone	52 (19.5)	101 (37.6)
Total	267	269

3 Models for longitudinal ordinal data

3.1 The proportional odds model

Consider a sample of N subjects and let Y be an ordered variable with K categories assessed on T occasions in each subject. Then, let Y_{ij} denote the assessment of the ordinal variable Y for the i th subject ($i = 1, \dots, N$) at the j th occasion ($j = 1, \dots, T$). Hence, $\mathbf{Y}_i =$

$(Y_{i1}, \dots, Y_{iT})'$ is the vector of the repeated assessments of the i th subject and $\mathbf{Y}_j = (Y_{1j}, \dots, Y_{Nj})'$ is the vector of responses at the j th occasion. Associated with each subject, there is a $p \times 1$ vector of covariates, say \mathbf{x}_{ij} , measured at time j . Hence, let $\mathbf{X}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{iT})'$ denote the $T \times p$ design matrix of the i th subject. Covariates typically include time of measurement, age, gender, treatment group, and so on.

The ordinal nature of the outcome variable may be accounted for by considering the cumulative probabilities $Pr(Y_{ij} \leq k), k = 1, \dots, K$. The cumulative proportional odds model is a popular choice to relate the marginal probabilities of Y to the covariate vector \mathbf{x} [10]. Specifically,

$$\text{logit}[Pr(Y_{ij} \leq k | \mathbf{x}_{ij})] = \beta_{0k} + \mathbf{x}_{ij}'\boldsymbol{\beta} \quad (1)$$

where $\boldsymbol{\beta}_0 = (\beta_{01}, \dots, \beta_{0,K-1})'$ is the vector of the intercept parameters and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ the vector of coefficients ($i = 1, \dots, N; j = 1, \dots, T; k = 1, \dots, K - 1$). Under the proportional odds assumption, $\boldsymbol{\beta}$ does not depend on k .

3.2 Generalized estimating equations

Estimation of the regression coefficients of marginal models can be approached by likelihood-based or non-likelihood-based methods. One difficulty present with likelihood models resides in the complexity of the relationship between the parameters of the model and the joint probabilities that define the likelihood. Alternative solutions to likelihood-based analysis have been explored, in particular the generalized estimating equations (GEE), quite popular for the analysis of non-Gaussian correlated data. This approach circumvents the specification of the joint distribution of the repeated responses by means of a ‘working’ correlation matrix and only the marginal distributions are

specified. Since the proportional odds model is not part of the regular generalized linear model family, some transformations are required before applying the GEE method.

Following Lipsitz *et al.* [11], a $(K - 1)$ -dimensional expanded vector of binary responses has to be created for each subject at each occasion, $\mathbf{Y}_{ij}^* = (Y_{i1j}^*, \dots, Y_{i,(K-1),j}^*)'$ where $Y_{ikj}^* = 1$ if $Y_{ij} = k$ and 0 otherwise. Now,

$$\text{logit}[\Pr(Y_{ij} \leq k | \mathbf{x}_{ij})] = \text{logit}[\Pr(Y_{ikj}^* = 1 | \mathbf{x}_{ij})], \quad k = 1, \dots, K - 1 \quad (2)$$

Since the logistic regression model is a member of the generalized linear model family, the GEE method applies and consistent estimates of the regression parameters can be obtained by solving the estimating equations

$$\sum_{i=1}^N \frac{\partial \boldsymbol{\pi}_i'}{\partial \boldsymbol{\beta}} \mathbf{V}_i^{-1} (\mathbf{Y}_i^* - \boldsymbol{\pi}_i) = \mathbf{0} \quad (3)$$

where $\mathbf{Y}_i^* = (\mathbf{Y}_{i1}^*, \dots, \mathbf{Y}_{iT}^*)'$, $\boldsymbol{\pi}_i = E(\mathbf{Y}_i^*)$, $\mathbf{V}_i = \mathbf{A}_i^{1/2} \mathbf{R}_i \mathbf{A}_i^{1/2}$ with \mathbf{A}_i the diagonal matrix of the variance of the elements of \mathbf{Y}_i^* , and $\boldsymbol{\beta}$ the expanded vector of intercepts and regression coefficients. The matrix \mathbf{R}_i is the ‘working’ correlation matrix that expresses the dependence among repeated observations over the subjects ranging from independence to exchangeable, banded, or unstructured.

4 Missingness

In line with the notation introduced previously, consider the missing data indicators, R_{ij} , defined as follows:

$$R_{ij} = \begin{cases} 1 & \text{if } Y_{ij} \text{ is observed,} \\ 0 & \text{otherwise,} \end{cases}$$

and let $\mathbf{R}_i = (R_{i1}, \dots, R_{iT})'$ the indicator vector corresponding to $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iT})'$.

Now \mathbf{Y}_i can be split into two subvectors $(\mathbf{Y}_i^o, \mathbf{Y}_i^m)$ where \mathbf{Y}_i^o refers to the observed component of \mathbf{Y}_i and \mathbf{Y}_i^m refers to the missing component part.

When missing data occur, we are concerned with the distribution of the measurement process together with the missing-data process. Little and Rubin [12, 13, 14] identified two broad classes of joint models: the selection model and the pattern-mixture model. In the selection model, the joint distribution $(\mathbf{Y}_i, \mathbf{R}_i)$ is split into the marginal distribution of the measurement and the distribution of the missingness process conditional on the measurement \mathbf{Y}_i . By contrast, the pattern-mixture model specifies the marginal distribution of \mathbf{R}_i and the conditional distribution of \mathbf{Y}_i given \mathbf{R}_i . Here we shall focus on the selection model approach in which Rubin [4] and Little and Rubin [12] made essential distinctions between the processes responsible for the missingness: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). The determination of the mechanism responsible for missing data has a decisive implication on the choice of the statistical method used to analyze the data. Under the MCAR mechanism, the probability of an observation being missing is independent of both \mathbf{Y}^o and \mathbf{Y}^m . Under the MAR mechanism, the probability of an observation being missing is independent of \mathbf{Y}^m given \mathbf{Y}^o . When neither MCAR nor MAR holds, the missingness mechanism is said to be MNAR, whence the probability of an observation being missing depends on \mathbf{Y}^m .

Liang and Zeger [15] pointed out that GEE are only valid under the restrictive assumption that the data are missing completely at random (MCAR). Alternative methods were investigated to allow the analysis of data under less strict missingness assumptions.

Robins *et al.* [16, 17] developed an extension of the GEE, known as the weighted generalized estimating equations (WGEE), that provide consistent estimates of the

regression parameters even under the MAR assumption. With their method, each subject's measurements is weighted in the GEE by the inverse probability of dropping out at that time point. Another alternative to analyze the data under the MAR assumption is multiple imputation based on GEE (MI-GEE). In this approach, missing values are imputed several times [4, 18] and the resulting completed datasets are analyzed using standard GEE methods. Using Rubin's rules, the final results obtained from the completed datasets are combined into a single inference. In the context of longitudinal binary data, Beunckens *et al.* [19] showed by simulations that, in spite of the asymptotic unbiasedness of WGEE, the combination of GEE and multiple imputation is both less biased and more accurate in small to moderate sample sizes which typically arise in clinical trials. In this paper, focus will be on MI-GEE methods.

5 Multiple imputation

5.1 Theoretical framework

The idea behind multiple imputation is to replace each missing value with a set of $M > 1$ plausible values drawn from the conditional distribution of the missing data given the observed data. This conditional distribution represents the uncertainty about the right value to impute in the sense that the set of M imputed values properly represents the information about the missing value that is contained in the observed data.

Using the notation introduced in previous sections, let $\boldsymbol{\theta}$ represents the parameter vector of the distribution of the response $\mathbf{Y}_i = (\mathbf{Y}_i^o, \mathbf{Y}_i^m)$. Note that $\boldsymbol{\theta}$ may differ from the parameters $\boldsymbol{\beta}$ of the substantive model. The observed data \mathbf{Y}^o will be used to estimate the conditional distribution of \mathbf{Y}^m given \mathbf{Y}^o , $f(\mathbf{Y}^m | \mathbf{Y}^o, \boldsymbol{\theta})$. If $\boldsymbol{\theta}$ is known, the values for \mathbf{Y}^m

can be drawn from $f(\mathbf{Y}^m|\mathbf{Y}^o, \boldsymbol{\theta})$. For $\boldsymbol{\theta}$ unknown, an estimate is obtained from the data, say $\hat{\boldsymbol{\theta}}$; then missing values will be imputed using $f(\mathbf{Y}^m|\mathbf{Y}^o, \hat{\boldsymbol{\theta}})$. Frequentists incorporate uncertainty in $\boldsymbol{\theta}$ by using bootstrap or other methods. A Bayesian prior distribution for $\boldsymbol{\theta}$ can also be chosen. Given this distribution, a draw $\boldsymbol{\theta}^*$ is generated and now values for \mathbf{Y}^m can be drawn from $f(\mathbf{Y}^m|\mathbf{Y}^o, \boldsymbol{\theta}^*)$. These two steps for the construction of the imputed data are the first phase of MI. Then the substantive model is applied to each of the M completed data $(\mathbf{Y}_i^o, \mathbf{Y}_i^{m*})$. Let $\hat{\boldsymbol{\beta}}_m$ and $\hat{\mathbf{U}}_m$ be the vector of estimates and the corresponding variance-covariance matrix for the m^{th} imputed data set ($m = 1, \dots, M$), respectively. The last step of MI is the combination of the M results. The MI point estimate for $\boldsymbol{\beta}$ is simply the average of the M complete-data point estimates [4, 5],

$$\hat{\boldsymbol{\beta}}^* = \frac{1}{M} \sum_{m=1}^M \hat{\boldsymbol{\beta}}_m$$

A measure of the precision of $\hat{\boldsymbol{\beta}}^*$ is obtained by Rubin's variance formula [4] which combines the within- and the between-imputation variability. Define \mathbf{W} , the within-imputation variance, as the average of the M within imputation variance estimates $\hat{\mathbf{U}}_m$,

$$\mathbf{W} = \frac{1}{M} \sum_{m=1}^M \hat{\mathbf{U}}_m$$

and \mathbf{B} , the between-imputation variance, measuring the variability across the imputed values,

$$\mathbf{B} = \frac{1}{M-1} \sum_{m=1}^M (\hat{\boldsymbol{\beta}}_m - \hat{\boldsymbol{\beta}}^*)(\hat{\boldsymbol{\beta}}_m - \hat{\boldsymbol{\beta}}^*)'$$

Then, the variance estimate associated with $\hat{\boldsymbol{\beta}}^*$ is the total variance

$$\mathbf{T} = \mathbf{W} + \left(1 + \frac{1}{M}\right) \mathbf{B}$$

where $(1 + \frac{1}{M})$ is a correction factor for the finite number of imputations.

5.2 MNI method

In Bayesian inference, information about unknown parameters is expressed in the form of posterior probability distributions computed using Bayes' theorem. In this context, Markov Chain Monte Carlo methods (MCMC) have been considered to explore and simulate the entire joint posterior distribution of the unknown quantities through the use of Markov chains.

Assuming that data arise from a multivariate normal distribution, Schafer [5] developed a method based on an MCMC process for generating proper imputations that accounts for between imputation variability, the MNI approach. This approach, based on the algorithm of data augmentation [20], is a procedure that iterates between an imputation step (I-step) and a posterior step (P-step). In the I-step, given starting values for the mean and the covariance matrix, i.e. given starting values for $\boldsymbol{\theta}$, values for missing data \mathbf{Y}^m are simulated by randomly drawing a value from the conditional multivariate normal distribution of \mathbf{Y}^m given \mathbf{Y}^o , $f(\mathbf{Y}^m|\mathbf{Y}^o, \boldsymbol{\theta})$. After the first iteration, new values for $\boldsymbol{\theta}$ are drawn from its posterior distribution. Both steps are iterated, which creates a Markov chain $(\mathbf{Y}_{(1)}^m, \boldsymbol{\theta}_{(1)}), (\mathbf{Y}_{(2)}^m, \boldsymbol{\theta}_{(2)}), \dots$ where each step depends on the previous one, introducing dependency across the steps. The two steps are then iterated long enough until the distribution becomes stationary. Imputations from the last iteration are used to impute the missing values of the dataset. More detail about this procedure can be found in [5].

When proceeding this way for an ordinal outcome, the imputed values obtained are no longer integer values and need then to be rounded off to the nearest integer (category) or to the nearest plausible value. However, in the binary case, it was demonstrated that rounding is not recommended because the rounded imputed values may provide biased

parameter estimates [21, 22, 23]. In situations like ours, where one is concerned with missing values for the outcome variable, unrounded values are physically not plausible. So, the rounding phase is unavoidable before application of the substantive model (e.g. GEE with proportional odds model).

5.3 FCS based on ordinal imputation model

The adaptation of the ordinal imputation model (OIM) to arbitrary missingness pattern appears as an alternative to the MNI approach. To impute missing data for an ordinal outcome, one has to impose a probability model on the complete data. In the presence of an ordinal outcome variable, a proportional odds model will be considered to link the ordinal outcome to a set of q covariates. The FCS with an ordinal imputation model is based on the Gibbs sampling algorithm; that is random draws from the multivariate distribution of interest, $f(\mathbf{Y}^m|\mathbf{Y}^o, \boldsymbol{\theta})$, is obtained by iteratively drawing from the conditional distribution of each outcome assessment. This imputation process is composed of two steps, a filled-in step and an imputation step.

Filled-in step

In this step, all missing value, \mathbf{Y}^m , are filled-in using an arbitrary method. Let $\mathbf{Y}^{(0)} = (\mathbf{Y}_1^{(0)}, \dots, \mathbf{Y}_T^{(0)})$ where $\mathbf{Y}_j^{(0)} = (\mathbf{Y}_j^o, \mathbf{Y}_{j*}^{(0)})$ with \mathbf{Y}_j^o the observed part of the j th assessment of the ordinal outcome Y and $\mathbf{Y}_{j*}^{(0)}$ its filled-in part. $\mathbf{Y}^{(0)}$ will serve as initial starting values for the imputation step.

Imputation step

In this second step, the previously filled-in elements of $\mathbf{Y}_{j*}^{(0)}$ are imputed using the specified conditional distribution, $f(\mathbf{Y}_j^m|\mathbf{Y}_j^o, \boldsymbol{\theta}_j)$. These imputations are made in turn for all \mathbf{Y}_j^m

($j = 1, \dots, T$). In order to obtain imputed values that are independent of the starting values, $\mathbf{Y}^{(0)}$, the cycling imputation through all \mathbf{Y}_j^m ($j = 1, \dots, T$) is repeated several times. The imputations above will be based on the following proportional odds model,

$$\text{logit}[Pr(Y_{ij}^m \leq k) | \mathbf{x}_{ij}^*] = \theta_{jk} + \mathbf{x}_{ij}^{\prime*} \boldsymbol{\theta}_{xj}, \quad (4)$$

where the covariates typically include those of the substantive model \mathbf{X}_{ij} , possible auxiliary covariates \mathbf{A}_{ij} , and the other outcomes $\mathbf{Y}_{-j} = (\mathbf{Y}_1, \dots, \mathbf{Y}_{(j-1)}, \mathbf{Y}_{(j+1)}, \dots, \mathbf{Y}_T)$. To realize proper imputation [4], uncertainty about $\boldsymbol{\theta}_j = (\theta_{jk}, \boldsymbol{\theta}_{xj})$ has to be accounted for. For this purpose, a value for $\boldsymbol{\theta}_j$ is drawn from an appropriate posterior distribution about $\boldsymbol{\theta}_j$ conditionally on the most recently imputed data. One way of proceeding is known as the "Normal approximation draw" method. This method is correct for linear regression [4] but is near far a reasonable approximation for situation involving categorical regression. Nevertheless, it is a common practice, supported by the law of large-sample, to use this Normal approximation [4]. To correct for possible misleading association that could have been introduced in the filled-in step, the proportional odds model is fitted on the part of the dataset with observed observation for the j th assessment, \mathbf{Y}_j^o , which might contain observations with imputed values for the other assessments, \mathbf{Y}_{-j} .

Based on these considerations, the t th iteration of the imputation step goes as follows,

$\mathbf{Y}_1^{(t)}$: 1. Fit the proportional odds model (4) on the part of the dataset for which \mathbf{Y}_1 is fully observed and draw new values for $\hat{\boldsymbol{\theta}}_1$ using

$$\boldsymbol{\theta}_{1*} = \hat{\boldsymbol{\theta}}_1 + \mathbf{V}_{hi}' \mathbf{Z},$$

where \mathbf{V}_{hi}' is the upper triangular matrix of the Cholesky decomposition,

$\mathbf{V}_i = \mathbf{V}_{hi}' \mathbf{V}_{hi}$ of the covariance matrix of $\hat{\boldsymbol{\theta}}_1$ and \mathbf{Z} is a $(K - 1) + q$ vector of

independent random normal variates.

2. For each element of \mathbf{Y}_1^m compute

$$P[Y_{i1}^m = k | \boldsymbol{\theta}_{1*}, \mathbf{Y}_1^o, \mathbf{Y}_2^{(t-1)}, \dots, \mathbf{Y}_T^{(t-1)}, \mathbf{x}_{i1}, \mathbf{A}_{i1}] \text{ from equation (4).}$$

3. For each element of \mathbf{Y}_1^m draw a random variate from a multinomial distribution with probabilities derived in step 2.

...

...

...

- $\mathbf{Y}_T^{(t)}$:
1. Fit the proportional odds model (4) on the part of the dataset for which \mathbf{Y}_T is fully observed and draw new values for $\hat{\boldsymbol{\theta}}_T$ using

$$\boldsymbol{\theta}_{T*} = \hat{\boldsymbol{\theta}}_T + \mathbf{V}_{hi}' \mathbf{Z},$$

where \mathbf{V}_{hi}' is the upper triangular matrix of the Cholesky decomposition,

$\mathbf{V}_i = \mathbf{V}_{hi}' \mathbf{V}_{hi}$ of the covariance matrix of $\hat{\boldsymbol{\theta}}_T$ and \mathbf{Z} is a $(K - 1) + q$ vector of independent random normal variates.

2. For each element of \mathbf{Y}_T^m compute $P[Y_{iT}^m = k | \boldsymbol{\theta}_{T*}, \mathbf{Y}_1^t, \mathbf{Y}_2^{(t)}, \dots, \mathbf{Y}_T^o, \mathbf{x}_{i1}, \mathbf{A}_{i1}]$ from equation (4).
3. For each element of \mathbf{Y}_T^m draw a random variate from a multinomial distribution with probabilities derived in step 2.

The previous cyclic iteration process is repeated several times, usually between 10-20

[25, 24], until stabilization of the results. As within the Gibbs sampling algorithm,

convergence is influenced by the choice of the initial values, $\mathbf{Y}^{(0)}$. In the filled-in step, we then replace the missing values using an ordinal logistic regression sequentially by order of assessment.

6 SIMULATION STUDY

To assess the performance of both imputation methods (MNI and FCS OIM), we conducted a large simulation study as described hereafter.

6.1 Longitudinal ordinal data-generating model

Correlated ordinal responses were generated with the SAS macro developed by Ibrahim [26] and based on Lee's algorithm [27]. The basic measurement model utilized in this study includes as covariates a binary group effect ($X = 0$ or 1), an assessment time (T) and an interaction term between group and time, so that the proportional odds model (Eq. 1) is written as:

$$\text{logit}[\Pr(Y_{ij} \leq k | x_i, t_j)] = \beta_{0k} + \beta_x x_i + \beta_t t_j + \beta_{tx} x_i t_j. \quad (5)$$

($i = 1, \dots, N$; $j = 1, \dots, T$; $k = 1, \dots, K - 1$). An exchangeable correlation structure was considered.

6.2 Missing data generating mechanisms

The mechanism used to generate MAR missingness data is based on the following binary logistic regression model:

$$\text{logit}[\Pr(R_{ij} = 0 | x_i, Y_{i,(j-1)})] = \psi_0 + \psi_x x_i + \psi_{prev} Y_{i,(j-1)} \quad (6)$$

($i = 1, \dots, N$; $j = 1, \dots, T$; $k = 1, \dots, K - 1$). Thus, the probability to be missing at a certain time point j depends on the binary covariate X and the outcome value at the previous time point $Y_{i,(j-1)}$.

6.3 Simulation patterns

Theoretical values of the model parameters (see (Eq. 5)) considered in our simulations are given in Table 3 for a well-balanced and skewed distribution.

Table 3: Values of the model parameters used for generating longitudinal ordinal dataset (well-balanced and skewed distribution)

Distribution	K	β_{01}	β_{02}	β_{03}	β_{04}	β_{05}	β_{06}	β_x	β_t	β_{tx}
Well-balanced										
	2	-0.25	-	-	-	-	-	0.10	0.10	-0.15
	3	-0.71	0.66	-	-	-	-	0.10	0.10	-0.15
	4	-1.10	0.00	1.10	-	-	-	0.10	0.10	-0.15
	5	-1.39	-0.41	0.41	1.39	-	-	0.10	0.10	-0.15
	7	-1.79	-0.92	-0.29	0.29	0.92	1.79	0.10	0.10	-0.15
Skewed										
	2	1.00	-	-	-	-	-	0.80	0.10	-0.25
	3	-2.20	-0.85	-	-	-	-	0.80	0.10	-0.25
	4	-0.41	0.00	0.41	-	-	-	0.80	0.10	-0.25
	5	-0.85	-0.20	0.20	0.85	-	-	0.80	0.10	-0.25
	7	-1.39	-0.66	-0.16	0.16	0.66	1.39	0.80	0.10	-0.25

Three distinct sample sizes N were considered for the simulation: 100, 300 and 500, equally distributed between both groups. This covers small (50 subjects/arm) to large studies (250 subjects/arm). For the assessment time points T , two possibilities were envisaged corresponding to short ($T = 3$) or long ($T = 5$) longitudinal study. Note that for skewed data, only $T = 3$ was considered. The ordinal outcome variable Y covered several numbers of categories $K = 2, 3, 4, 5$ and 7, respectively. Finally, the population parameters of (Eq. 6) ($\psi_0, \psi_x, \psi_{prev}$) were chosen to yield a rate of missingness approximatively equal to 10%, 30% and 50%, respectively. The complete data case (0% missingness) was also considered. Thus, both imputation methods were assessed on 90 different combination patterns. For each pattern, $S = 500$ random samples were generated. The two MI methods (MNI and FCS OIM) were applied to impute missing data on the same

incomplete dataset allowing a paired comparison of the two approaches. A GEE model was then fitted to the resulting multiply imputed datasets. For each MI method, the number of multiple imputation was fixed to $M = 20$ [4, 28]. As the generation of the MAR missingness was based on the binary covariate X , the latter had to be included in the imputation model. In the GEE model, the same working correlation matrix as the one used in the generation data process was considered, that is an exchangeable correlation matrix. The MI based on MNI and on FCS OIM were carried out using the SAS MI procedure. The GEE SAS macro based on the extension of Lipsitz *et al.* method [11] and implemented by Williamson *et al.* [29] was used to analyze the imputed datasets. Finally, the SAS MIANALYZE procedure was used to pool the results obtained.

6.4 Evaluation criteria

For each simulation pattern, the relative bias $RB = \hat{\beta}/\beta$ expressed in percent was averaged over the $S = 500$ replicated datasets. Likewise, the mean square error was calculated as

$$MSE = Bias^2 + Var(\hat{\beta})$$

with $Var(\hat{\beta}) = \sum_{s=1}^S \frac{(\hat{\beta}_s - \bar{\hat{\beta}})^2}{(S-1)}$, $\bar{\hat{\beta}} = \sum_{s=1}^S \frac{\hat{\beta}_s}{S}$ and $Bias = \bar{\hat{\beta}} - \beta$.

The effect of the modeling parameters on RB was assessed by multiple regression analysis and so was the difference between RB obtained by MNI and FCS OIM, respectively. To account for the matching between both imputation methods, a generalized linear mixed model taking all modeling parameters as covariates was applied to the MSE derived after imputation. This statistical scheme was applied to both kinds of generated ordinal data, well-balanced and skewed distribution.

7 Results

The values of the relative bias (%) and the MSE calculated over the 500 replicate samples are detailed in Appendices for both imputation methods. For clarity, results for intercepts were omitted.

7.1 Well-balanced distributions

Relative bias

Table 4 reports the mean (\pm SD) of RB of each regression parameter derived under both imputation methods as well as their differences. Globally, underestimated values of the model parameters were found using the MNI method, while estimates derived with the FCS OIM method were almost unbiased. Although differences between the two imputation methods were highly significant ($p < 0.0001$) for all regression parameters, the RB difference was small ($3 - 8\%$).

When considering the results under the various simulation patterns, the following observations could be made. For the binary effect parameter, β_x , using the MNI method, the RB was unchanged for K and rate of missingness but it varied according to the number of time points ($p = 0.001$) and to N ($p = 0.019$). In fact, RB was lower in short term than in long term studies ($92.9 \pm 15.9\%$ vs $101.8 \pm 10.5\%$; $p = 0.001$) and it decreased from $100.1 \pm 18.8\%$ for $N=100$ to $92.1 \pm 9.27\%$ for $N = 500$. Nearly the same conclusions applied for the RB derived under the FCS OIM process. The RB remained unchanged with T and the rate of missingness but decreased with K ($p = 0.009$) and with N ($p = 0.022$). The RB for the time effect parameter, β_t , and for the interaction term, β_{tx} , behaved similarly under both MI methods. It significantly decreased with K ($p < 0.0001$), the rate of missingness ($p < 0.05$) and increased with the number of time

point ($p < 0.05$) but was unchanged for N . Overall, for each simulation pattern, better RB values were obtained under the FCS OIM approach.

Table 4: Relative bias (mean \pm SD) of the parameters of the substantive model after imputation of the ordinal outcome using MNI and FCS OIM methods.

Globally and according to the modeling parameters

	β_x				β_t				β_{tx}			
	MNI	FCS OIM	Diff		MNI	FCS OIM	Diff		MNI	FCS OIM	Diff	
Global	97.4 \pm 14.1	100.1 \pm 12.7	-2.78 \pm 3.92 < 0.0001		90.4 \pm 14.1	98.3 \pm 9.22	-7.90 \pm 6.34 < 0.0001		95.2 \pm 6.09	99.2 \pm 4.57	-4.02 \pm 2.53 < 0.0001	
K	2	93.7 \pm 13.6	99.1 \pm 12.0	-5.36 \pm 3.84	99.0 \pm 4.87	102.4 \pm 6.31	-3.40 \pm 2.27		96.4 \pm 3.71	100.1 \pm 2.65	-3.70 \pm 2.65	
	3	110.5 \pm 12.9	111.7 \pm 12.9	-1.12 \pm 1.75	97.1 \pm 6.30	102.4 \pm 4.71	-5.34 \pm 2.44		99.6 \pm 5.48	102.9 \pm 4.77	-3.30 \pm 1.63	
	4	93.9 \pm 14.4	98.0 \pm 12.0	-4.05 \pm 4.69	88.0 \pm 11.2	97.6 \pm 4.91	-9.58 \pm 6.75		94.4 \pm 7.10	99.2 \pm 4.82	-4.80 \pm 3.29	
	5	95.5 \pm 10.8	97.5 \pm 12.1	-2.01 \pm 2.20	90.5 \pm 8.55	100.3 \pm 4.99	-9.86 \pm 6.56		93.3 \pm 4.75	97.5 \pm 3.00	-4.14 \pm 2.53	
	7	93.2 \pm 11.8	94.6 \pm 7.68	-1.35 \pm 4.63	77.5 \pm 21.7	88.8 \pm 14.1	-11.3 \pm 7.92		92.3 \pm 6.46	96.4 \pm 4.52	-4.16 \pm 2.31	
		0.086	0.009	0.003	<0.0001	<0.0001	<0.0001		<0.0001	<0.0001	0.063	
T	3	92.9 \pm 15.9	97.8 \pm 14.3	-4.89 \pm 3.86	85.4 \pm 18.1	96.7 \pm 12.6	-11.3 \pm 7.01		92.4 \pm 7.09	97.8 \pm 5.75	-5.34 \pm 2.61	
	5	101.8 \pm 10.5	102.5 \pm 10.6	-0.66 \pm 2.65	95.4 \pm 4.72	99.9 \pm 3.01	-4.51 \pm 2.93		98.0 \pm 3.02	100.7 \pm 2.22	-2.70 \pm 1.61	
		0.001	0.064	<0.001	<0.0001	0.046	<0.0001		<0.0001	0.0005	<0.0001	
N	100	100.1 \pm 18.8	102.4 \pm 17.2	-2.27 \pm 3.97	92.1 \pm 14.7	99.5 \pm 10.1	-7.36 \pm 6.17		96.0 \pm 7.77	99.6 \pm 6.39	-3.58 \pm 2.40	
	300	99.9 \pm 11.4	102.9 \pm 10.4	-2.94 \pm 4.01	90.4 \pm 14.4	98.4 \pm 9.32	-7.99 \pm 6.50		95.7 \pm 5.32	99.9 \pm 3.50	-4.17 \pm 2.60	
	500	92.1 \pm 9.27	95.2 \pm 7.57	-3.12 \pm 3.87	88.6 \pm 13.4	97.0 \pm 8.37	-8.35 \pm 6.53		93.9 \pm 4.73	98.2 \pm 3.04	-4.30 \pm 2.61	
		0.019	0.022	0.29	0.15	0.20	0.23		0.043	0.16	0.013	
Missingness	10	100.1 \pm 1.6	101.9 \pm 13.4	-1.74 \pm 1.32	97.3 \pm 6.76	100.8 \pm 5.55	-3.51 \pm 2.66		98.7 \pm 4.31	100.4 \pm 4.06	-1.71 \pm 0.87	
	30	97.9 \pm 14.7	100.8 \pm 13.1	-2.95 \pm 3.75	90.9 \pm 11.4	99.4 \pm 7.41	-8.54 \pm 5.86		95.6 \pm 5.62	99.7 \pm 4.58	-4.18 \pm 1.90	
	50	94.1 \pm 13.9	97.8 \pm 11.6	-3.64 \pm 5.43	83.1 \pm 18.2	94.7 \pm 12.4	-11.7 \pm 6.86		91.4 \pm 5.99	97.5 \pm 4.69	-6.17 \pm 2.23	
		0.074	0.18	0.020	<0.0001	0.003	<0.0001		<0.0001	0.0045	<0.0001	

Mean square error

The mean square error (mean \pm SD) of each regression parameters under both imputation methods and their difference are given in Table 5. Globally, although results were significant, difference between MNI and FCS OIM were minute and not practically relevant. From this perspective, MNI and FCS OIM were similar.

As expected, under both imputation methods and for each model parameter, the MSE decreased significantly ($p < 0.0001$) with the sample size N . A decrease was also observed with T ($p < 0.0001$). The number of categories K and rate of missingness did not affect MSE.

Table 5: Mean square error (mean \pm SD) of the parameters of the substantive model after imputation of the ordinal outcome using MNI and OIM methods.
Globally and according to the modeling parameters

		β_x			β_t			β_{tx}		
		MNI	FCS OIM	Diff	MNI	FCS OIM	Diff	MNI	FCS OIM	Diff
Global		0.120 \pm 0.101	0.117 \pm 0.098	0.003 \pm 0.007 < 0.0001	0.010 \pm 0.011	0.010 \pm 0.011	-0.000 \pm 0.001 0.008	0.019 \pm 0.021	0.019 \pm 0.022	-0.000 \pm 0.001 0.038
K	2	0.139 \pm 0.119	0.141 \pm 0.121	-0.002 \pm 0.002	0.011 \pm 0.013	0.012 \pm 0.013	-0.001 \pm 0.001	0.022 \pm 0.025	0.023 \pm 0.026	-0.001 \pm 0.002
	3	0.116 \pm 0.100	0.113 \pm 0.096	0.002 \pm 0.005	0.010 \pm 0.011	0.010 \pm 0.012	-0.000 \pm 0.001	0.019 \pm 0.022	0.020 \pm 0.022	-0.000 \pm 0.000
	4	0.122 \pm 0.108	0.117 \pm 0.102	0.005 \pm 0.007	0.010 \pm 0.012	0.010 \pm 0.012	-0.000 \pm 0.001	0.020 \pm 0.024	0.020 \pm 0.024	-0.000 \pm 0.000
	5	0.112 \pm 0.096	0.108 \pm 0.091	0.005 \pm 0.007	0.009 \pm 0.010	0.009 \pm 0.011	-0.000 \pm 0.000	0.017 \pm 0.019	0.017 \pm 0.019	0.000 \pm 0.000
	7	0.111 \pm 0.089 0.12	0.105 \pm 0.081 0.031	0.007 \pm 0.010 < 0.0001	0.011 \pm 0.012 0.65	0.010 \pm 0.011 0.36	0.000 \pm 0.001 < 0.0001	0.017 \pm 0.018 0.16	0.017 \pm 0.018 0.09	0.000 \pm 0.001 < 0.0001
T	3	0.160 \pm 0.118	0.154 \pm 0.115	0.006 \pm 0.009	0.017 \pm 0.012	0.018 \pm 0.013	-0.000 \pm 0.001	0.032 \pm 0.024	0.032 \pm 0.024	-0.000 \pm 0.001
	5	0.081 \pm 0.058 < 0.0001	0.080 \pm 0.058 < 0.0001	0.001 \pm 0.003 0.0001	0.003 \pm 0.002 < 0.0001	0.003 \pm 0.002 < 0.0001	-0.000 \pm 0.000 0.016	0.006 \pm 0.004 < 0.0001	0.006 \pm 0.005 < 0.0001	-0.000 \pm 0.000 0.038
N	100	0.240 \pm 0.087	0.233 \pm 0.084	0.007 \pm 0.011	0.020 \pm 0.014	0.021 \pm 0.015	-0.0001 \pm 0.001	0.038 \pm 0.027	0.038 \pm 0.028	-0.000 \pm 0.001
	300	0.075 \pm 0.025	0.074 \pm 0.024	0.002 \pm 0.003	0.007 \pm 0.005	0.007 \pm 0.005	-0.000 \pm 0.000	0.012 \pm 0.008	0.012 \pm 0.008	-0.000 \pm 0.000
	500	0.045 \pm 0.015	0.044 \pm 0.015	0.001 \pm 0.002	0.004 \pm 0.003	0.004 \pm 0.003	0.000 \pm 0.000	0.007 \pm 0.005	0.007 \pm 0.005	-0.000 \pm 0.000
		< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.0002	< 0.0001	< 0.0001	0.16
Missingness	10	0.116 \pm 0.100	0.115 \pm 0.099	0.001 \pm 0.002	0.009 \pm 0.011	0.009 \pm 0.011	-0.000 \pm 0.000	0.018 \pm 0.021	0.018 \pm 0.021	-0.000 \pm 0.000
	30	0.120 \pm 0.101	0.117 \pm 0.099	0.003 \pm 0.005	0.010 \pm 0.011	0.010 \pm 0.011	-0.000 \pm 0.001	0.019 \pm 0.021	0.019 \pm 0.021	-0.000 \pm 0.001
	50	0.125 \pm 0.105 0.46	0.119 \pm 0.098 0.73	0.006 \pm 0.011 0.0008	0.011 \pm 0.012 0.18	0.012 \pm 0.013 0.13	-0.000 \pm 0.001 0.040	0.020 \pm 0.023 0.44	0.021 \pm 0.023 0.40	-0.000 \pm 0.001 0.15

7.2 Skewed distributions

As mentioned in the simulation plan, the impact of both imputation methods within the skewed ordinal data setting has been investigated in the context of a short term study, that is $T = 3$. Simulation results are summarized in the Appendices.

The overall RBs under both imputation methods are depicted in Figure 1 for each regression parameter. Globally, the MNI method overestimated the binary and the interaction term parameters of the model, while at the same time underestimated the time parameter β_t . As in the well-balanced setting, the OIM method yielded less biased estimates. The median RB difference between the two imputation methods ranged from 2% to 10%, with the worst results observed for the time parameter, β_t . In fact, the lowest RB value of β_t was equal to 52.6% and the highest RB value was equal to 205.6%; both extremes values were obtained under the MNI method. The extreme RBs under the OIM method presented the same but less marked behaviour; they were equal to 76.7% and 144.4%, respectively.

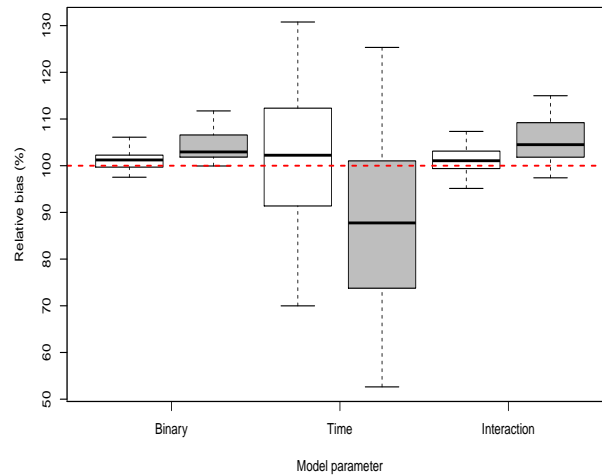


Figure 1: Global Relative bias (%) of the model parameters (β_x , β_t , β_{tx}) (MNI= shaded boxplot - OIM=empty boxplot)

The effect of the modeling parameters on the RB derived under both imputation methods

was found to be the same for K and N but not for the rate of missingness. As shown in Figure 2, under both multiple imputation methods, the RB varied according to K , especially for the time effect. While no association was found between RB and the rate of missingness for the OIM, Figure 3 shows that, except for the time effect, RB under MNI increased significantly with the rate of missingness (β_x : $p = 0.0003$, β_t : $p = 0.99$, β_{tx} : $p < 0.0001$). No relationship was observed between the RBs derived under both MI methods and the sample size, N .

The MSE of each regression parameter under both imputation methods and their differences are displayed in Table 6. Comparison of the MSE calculated in presence of skewed ordinal outcomes with those derived in well-balanced setting showed that MSE values were larger in presence of skewness. Contrary to the well-balanced setting, differences in the behaviors of the MSE were observed with respect to the modeling parameters, especially according to K .

As expected, under both imputation methods and for each model parameter, the MSE decreased significantly ($p < 0.0001$) with the sample size N . Contrary to the well-balanced setting, MSE values got lower as the number of categories K increased. However, these falls in the MSE behaved differently in the two MI methods for the binary and the interaction terms of the model. For the binary effect of the model, the difference in MSE increased with the number of categories of the ordinal outcome ($p < 0.0001$), while for the interaction term the MSE difference decreased ($p < 0.0001$). While the rate of missingness did not affect MSE; the difference in MSE between the two MI methods increased with the rate of missingness.

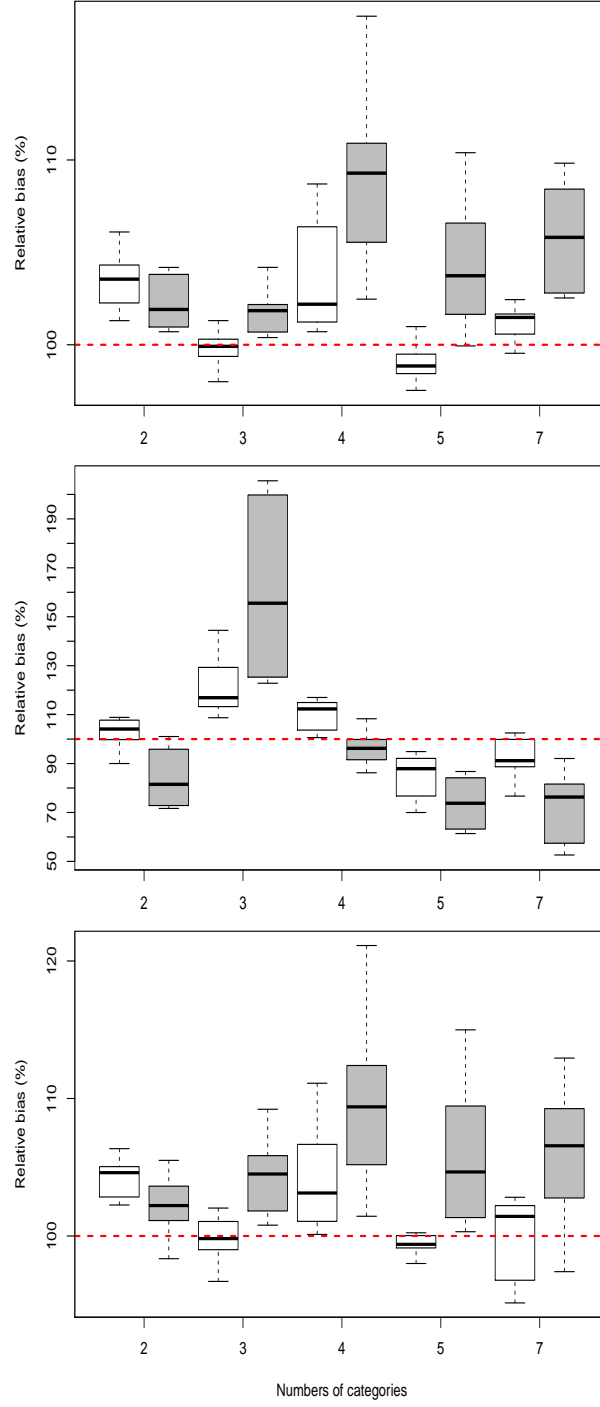


Figure 2: Relative bias (%) of the model parameters (β_x , β_t , β_{tx}) according to K the number of categories of the ordinal outcome (MNI= shaded boxplot - OIM=empty boxplot)

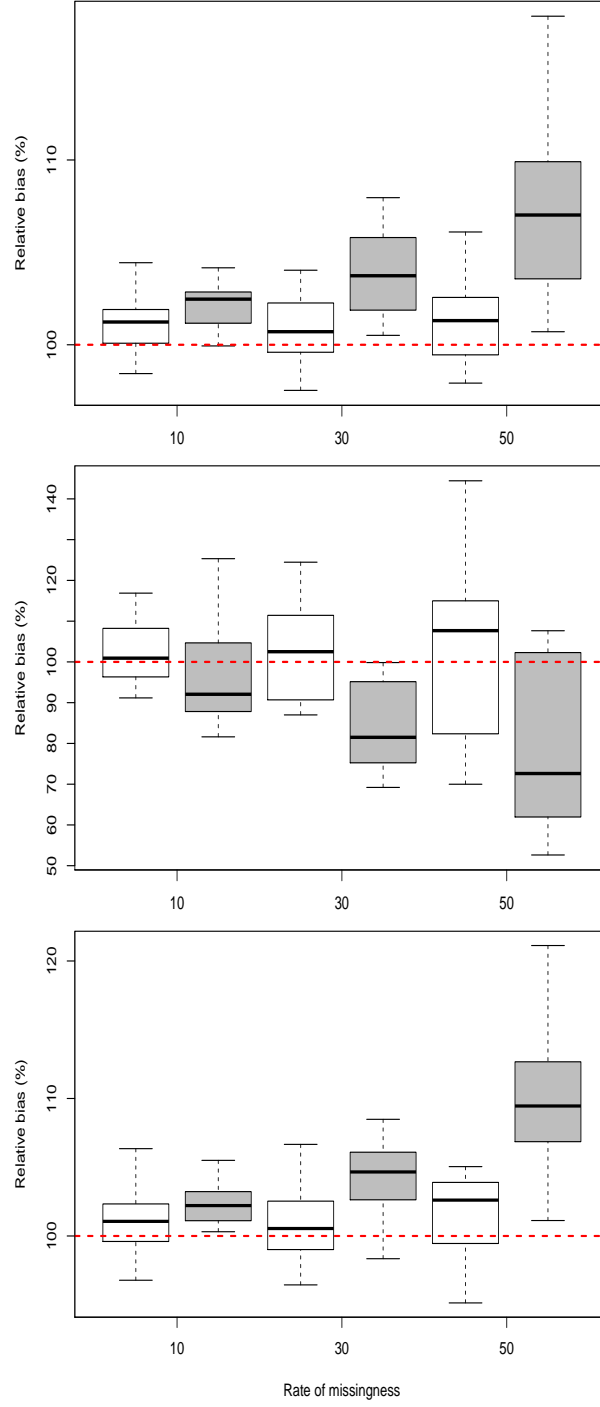


Figure 3: Relative bias (%) of the model parameters (β_x , β_t , β_{tx}) according to the rate of missingness (MNI=shaded boxplot - OIM=empty boxplot)

Table 6: Mean square error (mean \pm SD) of the parameters of the substantive model after imputation of the ordinal outcome using MNI and OIM methods, globally and according to the modeling parameters (skewed distribution)

	β_x				β_t				β_{tx}			
	MNI	FCS	OIM	Diff	MNI	FCS	OIM	Diff	MNI	FCS	OIM	Diff
Global	0.203 \pm 0.170	0.198 \pm 0.172	0.005 \pm 0.013	0.005 \pm 0.013 < 0.0001	0.021 \pm 0.017	0.021 \pm 0.017	0.000 \pm 0.002	0.000 \pm 0.002 0.84	0.041 \pm 0.035	0.042 \pm 0.037	-0.001 \pm 0.003	< 0.0001
K	2	0.339 \pm 0.264	0.350 \pm 0.271	-0.011 \pm 0.011	0.034 \pm 0.026	0.035 \pm 0.026	-0.001 \pm 0.001	-0.001 \pm 0.001	0.070 \pm 0.054	0.075 \pm 0.057	-0.005 \pm 0.005	
	3	0.180 \pm 0.132	0.179 \pm 0.131	0.001 \pm 0.002	0.024 \pm 0.015	0.021 \pm 0.014	0.003 \pm 0.004	0.003 \pm 0.004	0.036 \pm 0.026	0.038 \pm 0.028	-0.002 \pm 0.002	
	4	0.165 \pm 0.119	0.153 \pm 0.111	0.012 \pm 0.011	0.016 \pm 0.012	0.017 \pm 0.013	-0.001 \pm 0.001	-0.001 \pm 0.001	0.032 \pm 0.025	0.033 \pm 0.025	-0.001 \pm 0.001	
	5	0.174 \pm 0.135	0.165 \pm 0.129	0.008 \pm 0.008	0.017 \pm 0.013	0.017 \pm 0.014	-0.000 \pm 0.001	-0.000 \pm 0.001	0.034 \pm 0.027	0.035 \pm 0.027	-0.001 \pm 0.001	
	7	0.159 \pm 0.115	0.145 \pm 0.104	0.014 \pm 0.015	0.015 \pm 0.011	0.015 \pm 0.011	0.000 \pm 0.001	0.000 \pm 0.001	0.030 \pm 0.023	0.030 \pm 0.023	0.000 \pm 0.001	
		0.0015	0.0005	< 0.0001	< 0.0001	0.0001	0.87	0.87	0.0007	0.0003	< 0.0001	
N	100	0.404 \pm 0.148	0.395 \pm 0.164	0.009 \pm 0.021	0.041 \pm 0.015	0.041 \pm 0.015	-0.000 \pm 0.003	-0.000 \pm 0.003	0.081 \pm 0.032	0.084 \pm 0.035	-0.003 \pm 0.004	
	300	0.130 \pm 0.044	0.127 \pm 0.050	0.003 \pm 0.008	0.014 \pm 0.005	0.013 \pm 0.005	0.000 \pm 0.002	0.000 \pm 0.002	0.026 \pm 0.010	0.027 \pm 0.012	-0.001 \pm 0.002	
	500	0.076 \pm 0.022	0.074 \pm 0.026	0.002 \pm 0.005	0.009 \pm 0.004	0.008 \pm 0.003	0.001 \pm 0.002	0.001 \pm 0.002	0.015 \pm 0.006	0.016 \pm 0.007	-0.001 \pm 0.001	
		< 0.0001	< 0.0001	0.070	< 0.0001	< 0.0001	0.23	0.23	< 0.0001	< 0.0001	0.007	
Missingness	10	0.196 \pm 0.171	0.194 \pm 0.172	0.002 \pm 0.004	0.019 \pm 0.015	0.019 \pm 0.016	-0.000 \pm 0.000	-0.000 \pm 0.000	0.039 \pm 0.034	0.039 \pm 0.035	-0.000 \pm 0.001	
	30	0.201 \pm 0.171	0.196 \pm 0.173	0.005 \pm 0.010	0.021 \pm 0.017	0.021 \pm 0.017	0.000 \pm 0.001	0.000 \pm 0.001	0.040 \pm 0.035	0.041 \pm 0.036	-0.002 \pm 0.002	
	50	0.213 \pm 0.179	0.205 \pm 0.184	0.008 \pm 0.021	0.024 \pm 0.019	0.023 \pm 0.020	0.001 \pm 0.004	0.001 \pm 0.004	0.043 \pm 0.038	0.046 \pm 0.041	-0.003 \pm 0.004	
		0.62	0.77	0.094	0.11	0.19	0.32	0.32	0.50	0.36	0.007	

8 QoL data EXAMPLE

We applied both imputation methods on the QoL data (see Table 7).

Table 7: Results of the MI-GEE (proportional odds model) when using MNI and FCS OIM as multiple imputation method

Parameter	MNI		FCS - OIM	
	Estimate (SE)	P-value	Estimate (SE)	P-value
β_{01}	1.41 (0.17)	< 0.0001	1.46 (0.15)	< 0.0001
β_{02}	3.59 (0.21)	< 0.0001	2.94 (0.21)	< 0.0001
T_1	-0.36 (0.22)	0.11	-0.097 (0.20)	0.62
T_2	-0.73 (0.22)	0.001	-0.52 (0.22)	0.021
T_3	-0.92 (0.24)	0.0001	-0.43 (0.33)	0.20
T_4	-0.70 (0.35)	0.054	0.10 (0.36)	0.77
$TRT \times T_0$	0.21 (0.27)	0.44	0.26 (0.26)	0.32
$TRT \times T_1$	-0.52 (0.22)	0.017	-0.69 (0.23)	0.003
$TRT \times T_2$	-0.12 (0.22)	0.59	-0.23 (0.23)	0.32
$TRT \times T_3$	-0.26 (0.26)	0.32	-0.47 (0.37)	0.21
$TRT \times T_4$	0.01 (0.34)	0.97	-0.47 (0.42)	0.27

TRT is treatment (0 = RT, 1 = RT+TMZ); T0 = Baseline; T1 = During RT; T2 = After RT; T3 = FU1; T4 = FU2

Results derived under the MNI method showed that AP was more severe during RT ($p = 0.001$) and after RT ($p = 0.0001$) than at baseline. Moreover, severe AP affected more RT + TMZ patients than RT patients ($TRT \times T_1$; $p = 0.017$) during treatment. When applying the FCS OIM approach, the time effect disappeared except after RT ($p = 0.021$). As for the MNI approach, the deleterious effect was significantly higher in RT + TMZ patients ($p = 0.0003$). The difference between the two MI methods is evidenced in Figure 4 where the probabilities of each category at each assessment time in both treatment arms are displayed for both MI approaches.

Increasing the number of imputations up to 100 to test the robustness of the results did not change the conclusions.

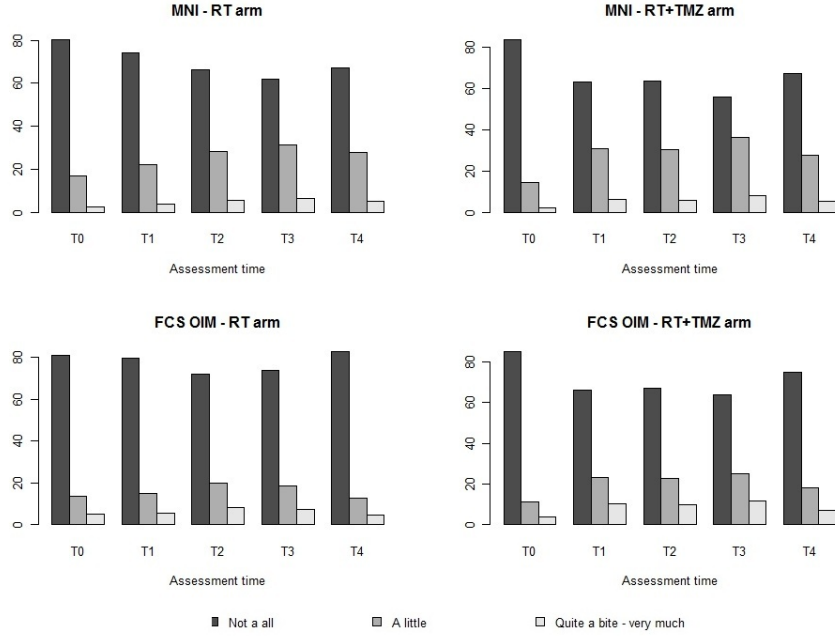


Figure 4: Distribution of appetite loss at each assessment time and in each treatment arm for both MI methods

9 DISCUSSION

Several studies have compared MNI and FCS imputation methods [25, 31, 30] but to the best of our knowledge, none have focused on longitudinal ordinal outcome data. This study was designed to compare the performance of the two methods, available in most statistical packages, in the context of longitudinal ordinal datasets with non-monotone missing values. The comparison was based on a comprehensive simulation plan covering a wide range of real life situations. Specifically, the parameters of the experimental design included the following parameters: number of categories of the ordinal outcome (K), number of time points (T), sample size (N) and rate of missingness (%) but also the form of the distribution (well-balanced or skewed) of the ordinal outcome data. Both MI methods were also applied on a real QoL dataset. The performance of the two MI

methods was appraised by the relative bias and the mean square error of the regression parameters of the model. The latter included a group effect and a time effect, as well as their interaction.

Within the well-balanced setting, the model parameters were slightly underestimated in the MNI approach as compared to the FCS OIM method which yielded almost unbiased estimates. Except for the binary term where effects were less marked, both imputation methods behaved similarly for each regression parameters. Under both MI methods, RB decreased with K and the rate of missingness, increased with the number of assessment time and was unchanged for the sample size (N). However, within each simulation pattern, RB values derived under the FCS OIM were slightly better than those derived under the MNI process. For all regression parameters, the MSE of both imputation methods were comparable.

For skewed data, application of the MNI process led to a marked overestimation of the regression coefficients of the binary and the interaction terms and an underestimation of the time coefficient. Overall, estimates derived under FCS OIM process were less biased. While, RB evolved differently according to K under both MI methods, it was only affected by the rate of missingness under MNI. In both distribution settings, estimation of the time effect coefficient was more biased than the other coefficients.

Although globally, simulations did not evidenced a large differences between the performance of the two MI methods, some simulation patterns were clearly against MNI. This was confirmed by the AP dataset where the ordinal outcome had $K=3$ categories, a skewed distribution and a large amount of missing data. Application of the two MI methods led to different conclusions, in particular for the time effect.

Within the longitudinal setting, Donneau *et al.* [6] previously showed that the OIM

method provides less biased results when imputing drop out cases than the MNI method. In comparison with those findings where the RB difference between the two imputation methods ranged from 9% to 16%, the difference between the MNI and the FCS OIM method found here was much lower (3% to 8%). As far as the MSE is concerned, the conclusions made for the non-monotone setting paralleled those found for the monotone setting.

Based on the results of this large simulation study and application to QoL dataset, salient conclusions may be drawn. Although theoretically unsuitable for ordinal data, the MNI method with rounding imputation to the nearest integer value globally provided better acceptable results than expected. However, as shown across the different simulation patterns, some situations were less favorable for MNI than for FCS OIM. This remark was reinforced by results of the QoL dataset where different conclusions applied according to the MI method used. Finally, as for the analysis model, the choice of the imputation method should be guided by the type of the data that needs to be imputed. Thus, it is advisable to impute missing ordinal data using suitable MI method.

References

1. Olschewski, M., Schulgen, G., Schumacher, M. and Altman, D.G. Quality of life assessment in clinical cancer research. *British journal of cancer* 1994. **70**: 1–5.
2. Carpenter, J.R., Kenward, M. G. Missing data in randomised controlled trials a practical guide Birmingham: National Institute for Health Research 2007. Available at www.missingdata.org.uk. **1–206**.
3. Rubin, D. Multiple imputation in sample surveys - a phenomenological bayesian

- approach to nonresponse. *Imputation and Editing of Faulty or Missing Survey Data* 1978. **1–32**.
4. Rubin, D. B. *Multiple imputations for nonresponse in survey* . Wiley: New York, 1987.
 5. Schafer, J. L. Multiple imputation for Nonresponse in Survey *Chapman & Hall*1997
 6. Donneau, A.F., Mauer, M., Molenberghs, G. and Albert, A. A simulation study comparing multiple imputation methods for incomplete longitudinal ordinal data. 2012 **submitted**
 7. Stupp R, Mason WP, van den Bent MJ et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine* 2005; **352**(10):987–996.
 8. Taphoorn MJ, Stupp R, Coens C et al. Health-related quality of life in patients with glioblastoma: a randomized controlled trial. *Lancet Oncology* 2005; **6**(12):937–944.
 9. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman DB, De Haes JCJM, Kaasa S, Klee M, Osoba D, Razavi D, Rofo P, Schraub S, Sneeuw K, Sullivan M, Takeda F. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute* 1993; **85**:365–376.
 10. McCullagh, P. Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society, Series B* 1980; **42**:109–142.

11. Lipsitz, SR., Kim, K., Zhao, L. Analysis of repeated categorical data using generalized estimating equations. *Statistics in Medicine* 1994; **13**(11):1149–1163.
12. Little, R. J. A., Rubin, D. B. *Statistical Analysis with Missing Data* . Wiley: New York, 1987.
13. Little, R. J. A. Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association* 1993. **88**:125–134.
14. Little, R. J. A. Modeling the drop-out mechanism in repeated measures studies. *Journal of the American Statistical Association* 1995. **90**:1112–1121.
15. Liang, K.-Y., Zeger, S. L. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**:13–22.
16. Robins, J. M., Rotnitzky, A. Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association* 1995; **90**:122–129.
17. Robins, J. M., Rotnitzky, A., Zhao, L. Analysis of semiparametric regression models with missing data. *Journal of the American Statistical Association* 1995; **90**:106–121.
18. Rubin, D. B. Inference and missing data. *Biometrika* 1976; **63**:581–592.
19. Beunckens, C., Sotto, C., Molenberghs, G. A simulation study comparing weighted estimating equations with multiple imputation based estimating equations for longitudinal binary data. *Computational Statistics and Data Analysis* 2008; **52**:1533–1548.

20. Tanner, M. A., Wong, W. H. The calculation of posterior distribution by data augmentation *Journal of American Statistical Association* 1987, **82**: 528–550.
21. Horton, N., Lipsitz, S., Parzen, M. A potential for bias when rounding in multiple imputation *The American Statistician* 2003, **57**:229–232.
22. Ake, C. Rounding after multiple imputation with non-binary categorical covariates. *Paper presented at SAS Users Group international 2005. Thirty annual conference, Philadelphia.*
23. Allison, P. Imputation of categorical variables with PROC MI *Paper presented at SAS Users Group international 2005. Thirty annual conference, Philadelphia*
24. White, Ian R. and Royston, Patrick and Wood, Angela M. Multiple imputation using chained equations: Issues and guidance for practice *Statistics in Medicine* 2011, **30**(4):377–399.
25. van Buuren, S. Multiple imputation of discrete and continuous data by full conditional specification *Statistical Methods in Medical Research* 2007 **16**:219–242.
26. Ibrahim, N., Suliadi, S. Generating correlated discrete ordinal data using R and SAS IML *Computer Methods and Programs in Biomedicine* 2011; **104**(3):122–132.
27. Lee, A. J. Some simple methods for generating correlated categorical variates *Computational Statistics and Data Analysis* 1997; **26**:133–148.
28. Graham, J. W., Olchowski, A. E., Gilreath, T. D. How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory *Prevention Science* 2007; **8**:206-213.

29. Williamson, J., Lipsitz, S., Kim, K. GEECAT and GEEGOR: computer programs for the analysis of correlated categorical response data. *Computer Methods and Programs in Biomedicine* 1999; **58**:25–34 .
30. Lee, K. and Carlin, J. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation *American Journal of Epidemiology* 2012 **71**:624–632.
31. Demirtas, H., Freels, S. and Yucel R. Plausibility of multivariate normality assumption when multiply imputing non-Gaussian continuous outcomes: a simulation assessment. *Journal of Statistical Computation and Simulation* 2008 **78**:69–84.

10 Appendices

Table 8: Simulation results for the MI-GEE based MNI and OIM methods (K = 2 - Well-balanced distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	87.7	0.389		90.1	0.393		88.7	0.372		95.2	0.376		83.9	0.360		97.0	0.366		83.9	0.360		97.0	0.366	
		β_t	104.2	0.037		106.0	0.038		107.0	0.037		111.9	0.040		109.4	0.039		116.6	0.043		109.4	0.039		116.6	0.043	
		β_{tx}	96.9	0.075		98.5	0.076		98.0	0.073		102.4	0.075		93.2	0.074		102.6	0.081		93.2	0.074		102.6	0.081	
3	300	β_x	93.8	0.104		95.9	0.104		89.7	0.102		96.5	0.103		88.3	0.103		99.1	0.107		88.3	0.103		99.1	0.107	
		β_t	99.1	0.012		100.4	0.012		99.2	0.012		103.2	0.013		105.6	0.012		111.8	0.013		105.6	0.012		111.8	0.013	
		β_{tx}	98.2	0.021		99.8	0.021		95.5	0.021		100.4	0.022		93.2	0.021		100.9	0.024		93.2	0.021		100.9	0.024	
3	500	β_x	75.5	0.071		79.1	0.071		72.9	0.066		79.9	0.067		72.0	0.067		84.4	0.068		72.0	0.067		84.4	0.068	
		β_t	94.8	0.007		96.6	0.008		94.1	0.007		98.1	0.008		102.6	0.008		110.9	0.009		102.6	0.008		110.9	0.009	
		β_{tx}	92.3	0.015		95.2	0.015		90.1	0.014		94.6	0.014		89.4	0.015		98.1	0.016		89.4	0.015		98.1	0.016	
5	100	β_x	112.6	0.186		113.7	0.187		118.4	0.195		119.8	0.196		113.5	0.195		120.5	0.201		113.5	0.195		120.5	0.201	
		β_t	97.5	0.007		98.1	0.007		98.7	0.008		100.7	0.008		97.7	0.008		101.3	0.009		97.7	0.008		101.3	0.009	
		β_{tx}	101.6	0.016		102.2	0.016		103.0	0.016		105.1	0.016		99.8	0.017		103.6	0.017		99.8	0.017		103.6	0.017	
5	300	β_x	105.1	0.063		106.3	0.63		102.8	0.063		105.5	0.065		103.4	0.064		110.6	0.067		103.4	0.064		110.6	0.067	
		β_t	97.6	0.002		98.4	0.003		96.9	0.003		99.1	0.003		95.6	0.003		100.3	0.003		95.6	0.003		100.3	0.003	
		β_{tx}	99.7	0.005		100.6	0.005		98.1	0.005		100.2	0.005		96.8	0.005		101.2	0.006		96.8	0.005		101.2	0.006	
5	500	β_x	96.4	0.035		98.0	0.036		93.0	0.035		96.2	0.035		88.9	0.035		95.5	0.037		88.9	0.035		95.5	0.037	
		β_t	96.3	0.001		97.3	0.001		94.5	0.001		96.8	0.001		91.5	0.001		95.8	0.001		91.5	0.001		95.8	0.001	
		β_{tx}	98.4	0.003		99.4	0.003		96.9	0.003		99.2	0.003		93.9	0.003		98.5	0.003		93.9	0.003		98.5	0.003	

Table 9: Simulation results for the MI-GEE based on MNI and OIM methods (K = 3 - Well-balanced distribution)

T	N	Parm	0%			10%			30%			50%			
			MNI		MSE	FCS OIM		MNI		FCS OIM		MNI		FCS OIM	
			RB(%)	MSE		RB(%)	MSE	RB(%)	MSE	RB(%)	MSE	RB(%)	MSE	RB(%)	MSE
3	100	β_x	137.2	0.298	140.8	0.296	131.0	0.315	131.0	0.307	118.6	0.327	115.8	0.308	
		β_t	108.5	0.031	112.8	0.032	100.9	0.032	107.5	0.033	89.8	0.036	96.5	0.038	
		β_{tx}	112.3	0.062	114.9	0.063	109.4	0.064	113.0	0.065	99.1	0.068	103.6	0.070	
3	300	β_x	112.9	0.091	115.2	0.090	104.1	0.096	106.7	0.093	102.1	0.097	100.8	0.093	
		β_t	101.9	0.010	106.1	0.010	93.3	0.011	102.3	0.011	88.8	0.011	96.3	0.012	
		β_{tx}	104.2	0.019	106.3	0.019	98.1	0.020	103.2	0.021	95.0	0.021	100.3	0.022	
3	500	β_x	97.4	0.053	99.4	0.052	99.6	0.054	100.7	0.053	92.6	0.058	92.9	0.055	
		β_t	95.9	0.006	100.2	0.006	91.0	0.007	99.6	0.007	86.3	0.007	94.6	0.008	
		β_{tx}	97.3	0.011	99.3	0.011	95.8	0.012	99.2	0.012	89.5	0.012	95.6	0.013	
5	100	β_x	122.0	0.147	122.5	0.146	117.9	0.150	116.7	0.150	118.4	0.154	120.0	0.152	
		β_t	105.4	0.007	106.6	0.007	102.9	0.007	106.8	0.007	98.5	0.007	104.7	0.007	
		β_{tx}	103.6	0.011	104.4	0.011	101.0	0.012	103.1	0.012	98.7	0.012	102.6	0.012	
5	300	β_x	115.1	0.049	115.4	0.049	114.0	0.049	117.1	0.049	116.4	0.050	118.3	0.051	
		β_t	103.5	0.002	105.0	0.002	100.0	0.002	104.1	0.002	94.6	0.002	101.0	0.002	
		β_{tx}	102.3	0.004	103.2	0.004	100.2	0.004	103.0	0.004	98.0	0.004	102.5	0.004	
5	500	β_x	98.9	0.030	99.8	0.030	97.9	0.031	99.6	0.030	93.6	0.031	97.1	0.031	
		β_t	99.6	0.001	101.2	0.001	96.0	0.001	100.3	0.001	90.2	0.001	97.5	0.001	
		β_{tx}	99.2	0.003	100.2	0.003	97.0	0.003	99.7	0.003	93.1	0.003	98.1	0.003	

Table 10: Simulation results for the MI-GEE based on MNI and OIM methods (K = 4 - Well-balanced distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	73.6	0.318		77.7	0.312		63.0	0.332		70.8	0.317		67.0	0.368		72.9	0.340		67.0	0.368		72.9	0.340	
		β_t	98.9	0.031		105.0	0.030		83.4	0.032		98.1	0.032		72.2	0.039		91.0	0.042		72.2	0.039		91.0	0.042	
		β_{tx}	90.5	0.064		93.2	0.064		82.7	0.066		89.0	0.065		79.8	0.079		88.3	0.079		79.8	0.079		88.3	0.079	
3	300	β_x	93.2	0.098		96.6	0.097		92.5	0.109		101.1	0.104		84.9	0.119		98.1	0.111		84.9	0.119		98.1	0.111	
		β_t	91.7	0.008		98.0	0.008		80.6	0.010		96.3	0.010		66.9	0.012		89.1	0.012		66.9	0.012		89.1	0.012	
		β_{tx}	95.3	0.018		98.0	0.018		93.2	0.020		100.7	0.020		85.0	0.024		96.3	0.024		85.0	0.024		96.3	0.024	
3	500	β_x	100.3	0.057		104.4	0.056		94.2	0.059		103.5	0.057		81.4	0.067		95.1	0.063		81.4	0.067		95.1	0.063	
		β_t	92.9	0.006		99.7	0.006		81.4	0.006		97.9	0.006		64.3	0.008		85.5	0.008		64.3	0.008		85.5	0.008	
		β_{tx}	99.1	0.011		102.2	0.011		93.4	0.012		101.3	0.012		85.3	0.014		96.6	0.015		85.3	0.014		96.6	0.015	
5	100	β_x	107.0	0.138		106.3	0.137		107.9	0.146		108.8	0.143		103.8	0.146		103.6	0.141		103.8	0.146		103.6	0.141	
		β_t	100.9	0.006		102.4	0.006		97.9	0.006		102.7	0.006		90.1	0.006		98.0	0.007		90.1	0.006		98.0	0.007	
		β_{tx}	103.4	0.011		104.3	0.011		102.1	0.011		104.5	0.011		97.5	0.011		101.7	0.011		97.5	0.011		101.7	0.011	
5	300	β_x	110.3	0.047		110.9	0.047		110.0	0.048		110.0	0.048		103.1	0.052		104.1	0.050		103.1	0.052		104.1	0.050	
		β_t	99.4	0.002		101.2	0.002		96.1	0.002		100.9	0.002		87.6	0.002		95.7	0.002		87.6	0.002		95.7	0.002	
		β_{tx}	102.5	0.004		103.5	0.004		100.4	0.004		103.3	0.004		95.9	0.004		100.8	0.004		95.9	0.004		100.8	0.004	
5	500	β_x	99.7	0.029		100.3	0.029		99.3	0.030		98.8	0.029		99.4	0.032		100.3	0.031		99.4	0.032		100.3	0.031	
		β_t	97.7	0.001		99.6	0.001		93.9	0.001		99.0	0.001		87.9	0.001		96.2	0.001		87.9	0.001		96.2	0.001	
		β_{tx}	99.8	0.002		100.9	0.002		98.0	0.002		100.8	0.002		95.0	0.003		100.0	0.003		95.0	0.003		100.0	0.003	

Table 11: Simulation results for the MI-GEE based on MNI and OIM methods (K = 5 - Well-balanced distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	105.1	0.289		107.1	0.284		104.4	0.293		108.9	0.281		89.7	0.305		108.9	0.281		89.7	0.305		95.2	0.276	
		β_t	99.4	0.028		105.4	0.028		87.6	0.030		102.0	0.031		76.1	0.032		102.0	0.031		76.1	0.032		97.0	0.034	
		β_{tx}	94.4	0.056		96.7	0.056		88.6	0.057		93.7	0.057		81.3	0.059		93.7	0.057		81.3	0.059		90.0	0.058	
3	300	β_x	118.8	0.083		121.5	0.081		113.6	0.087		120.1	0.085		105.2	0.091		120.1	0.085		105.2	0.091		105.0	0.083	
		β_t	103.3	0.009		109.5	0.009		90.9	0.010		108.3	0.011		76.0	0.012		108.3	0.011		76.0	0.012		94.3	0.012	
		β_{tx}	99.8	0.016		102.3	0.016		94.0	0.018		100.7	0.019		86.3	0.020		100.7	0.019		86.3	0.020		93.4	0.020	
3	500	β_x	103.7	0.049		106.6	0.049		100.4	0.052		106.4	0.050		97.2	0.056		106.4	0.050		97.2	0.056		98.2	0.051	
		β_t	98.8	0.005		105.1	0.005		87.7	0.006		105.4	0.006		74.3	0.007		105.4	0.006		74.3	0.007		92.9	0.007	
		β_{tx}	97.7	0.010		100.1	0.010		92.9	0.011		99.6	0.011		87.8	0.012		99.6	0.011		87.8	0.012		94.9	0.012	
5	100	β_x	87.5	0.155		88.3	0.152		83.4	0.169		83.6	0.164		85.4	0.172		83.6	0.164		85.4	0.172		86.3	0.161	
		β_t	94.3	0.005		95.6	0.005		89.9	0.006		94.1	0.006		85.8	0.006		94.1	0.006		85.8	0.006		94.7	0.007	
		β_{tx}	96.7	0.010		97.3	0.010		94.2	0.011		96.3	0.011		92.2	0.012		96.3	0.011		92.2	0.012		96.2	0.011	
5	300	β_x	88.0	0.044		88.8	0.044		88.8	0.047		89.8	0.046		85.6	0.050		89.8	0.046		85.6	0.050		85.6	0.048	
		β_t	99.3	0.002		100.7	0.002		95.7	0.002		101.7	0.002		88.1	0.002		101.7	0.002		88.1	0.002		99.0	0.002	
		β_{tx}	98.0	0.003		98.8	0.003		96.2	0.003		99.3	0.003		92.7	0.004		99.3	0.003		92.7	0.004		98.1	0.004	
5	500	β_x	88.1	0.026		88.3	0.026		87.2	0.028		87.7	0.027		86.8	0.029		87.7	0.027		86.8	0.029		87.6	0.028	
		β_t	98.8	0.001		100.2	0.001		94.2	0.001		100.0	0.001		88.4	0.001		100.0	0.001		88.4	0.001		100.1	0.001	
		β_{tx}	98.2	0.002		98.9	0.002		95.7	0.002		98.8	0.002		93.1	0.002		98.8	0.002		93.1	0.002		99.1	0.002	

Table 12: Simulation results for the MI-GEE based on MNI and OIM methods (K = 7 - Well-balanced distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	89.7	0.266		91.7	0.258		80.3	0.281		86.3	0.262		79.9	0.300		85.7	0.257		79.9	0.300		85.7	0.257	
		β_t	81.1	0.029		90.7	0.029		60.7	0.033		77.4	0.033		40.7	0.038		64.6	0.037		40.7	0.038		64.6	0.037	
		β_{tx}	95.0	0.051		98.3	0.051		87.5	0.053		94.3	0.052		84.0	0.060		89.0	0.057		84.0	0.060		89.0	0.057	
3	300	β_x	88.4	0.091		91.1	0.090		83.8	0.100		90.4	0.094		76.0	0.112		84.6	0.102		76.0	0.112		84.6	0.102	
		β_t	78.6	0.011		86.7	0.010		62.0	0.013		81.3	0.012		37.7	0.017		61.9	0.016		37.7	0.017		61.9	0.016	
		β_{tx}	92.2	0.018		95.0	0.018		86.7	0.019		93.1	0.019		80.9	0.023		89.6	0.022		80.9	0.023		89.6	0.022	
3	500	β_x	90.0	0.053		93.3	0.052		82.1	0.058		87.9	0.053		74.7	0.066		80.6	0.060		74.7	0.066		80.6	0.060	
		β_t	81.4	0.005		89.9	0.005		63.7	0.007		83.5	0.006		38.1	0.011		63.0	0.009		38.1	0.011		63.0	0.009	
		β_{tx}	92.6	0.010		95.7	0.010		85.8	0.011		92.3	0.010		80.3	0.013		87.5	0.013		80.3	0.013		87.5	0.013	
5	100	β_x	106.6	0.142		107.5	0.141		112.4	0.145		105.6	0.139		106.1	0.158		101.3	0.146		106.1	0.158		101.3	0.146	
		β_t	99.5	0.005		101.8	0.005		96.0	0.005		100.1	0.005		89.0	0.006		94.6	0.006		89.0	0.006		94.6	0.006	
		β_{tx}	100.1	0.010		101.3	0.010		99.1	0.011		100.4	0.011		94.6	0.012		97.3	0.012		94.6	0.012		97.3	0.012	
5	300	β_x	100.8	0.046		100.9	0.046		102.5	0.048		101.1	0.048		104.1	0.052		98.4	0.051		104.1	0.052		98.4	0.051	
		β_t	99.2	0.002		101.8	0.002		94.3	0.002		100.5	0.002		89.5	0.002		97.6	0.002		89.5	0.002		97.6	0.002	
		β_{tx}	99.5	0.004		100.8	0.004		97.5	0.004		100.9	0.004		95.3	0.004		100.0	0.004		95.3	0.004		100.0	0.004	
5	500	β_x	98.0	0.027		98.2	0.027		101.0	0.029		99.3	0.028		101.3	0.031		98.10	0.030		101.3	0.031		98.10	0.030	
		β_t	99.0	0.001		101.7	0.001		95.6	0.001		102.6	0.001		88.1	0.001		98.5	0.001		88.1	0.001		98.5	0.001	
		β_{tx}	98.8	0.002		100.2	0.002		96.8	0.002		100.3	0.002		93.7	0.002		99.4	0.002		93.7	0.002		99.4	0.002	

Table 13: Simulation results for the MI-GEE based on MNI and OIM methods ($K = 2$ - Skewed distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	104.2	0.680		104.4	0.685		101.0	0.673		101.3	0.687		104.2	0.706		106.1	0.744							
		β_t	95.8	0.059		99.8	0.059		74.2	0.066		90.0	0.066		72.8	0.077		98.8	0.081							
		β_{tx}	103.9	0.135		104.4	0.137		98.4	0.136		99.0	0.144		103.4	0.151		105.0	0.168							
3	300	β_x	103.8	0.214		104.3	0.216		102.9	0.210		104.0	0.217		101.3	0.217		103.6	0.233							
		β_t	101.0	0.020		107.7	0.020		87.8	0.021		107.9	0.021		71.7	0.023		107.7	0.024							
		β_{tx}	105.5	0.044		106.4	0.045		103.6	0.044		105.4	0.047		101.1	0.046		104.7	0.052							
3	500	β_x	101.9	0.120		102.3	0.121		100.8	0.118		102.1	0.122		100.7	0.114		102.9	0.123							
		β_t	97.8	0.012		104.1	0.012		81.5	0.013		102.2	0.013		72.6	0.014		108.9	0.015							
		β_{tx}	102.2	0.026		102.8	0.027		100.0	0.026		102.3	0.028		101.1	0.025		104.6	0.030							

Table 14: Simulation results for the MI-GEE based on MNI and OIM methods ($K = 3$ - Skewed distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	100.4	0.366		100.0	0.368		100.5	0.355		98.00	0.353		103.0	0.340		99.4	0.333							
		β_t	125.3	0.037		113.3	0.038		159.4	0.041		124.5	0.040		205.6	0.048		144.4	0.041							
		β_{tx}	100.8	0.073		99.4	0.074		101.7	0.069		96.70	0.072		107.9	0.070		101.1	0.075							
3	300	β_x	100.7	0.117		99.9	0.116		101.9	0.115		99.7	0.114		102.1	0.115		98.8	0.114							
		β_t	123.5	0.013		109.6	0.013		155.5	0.016		116.9	0.014		199.8	0.023		129.3	0.015							
		β_{tx}	101.8	0.024		99.8	0.024		104.5	0.024		99.0	0.025		105.8	0.024		97.7	0.026							
3	500	β_x	101.7	0.072		100.8	0.071		102.2	0.069		100.3	0.068		104.2	0.70		101.3	0.070							
		β_t	122.8	0.008		108.7	0.008		155.4	0.010		116.5	0.008		201.8	0.018		130.8	0.009							
		β_{tx}	103.6	0.014		101.4	0.014		105.3	0.013		100.5	0.014		109.2	0.013		102.0	0.015							

Table 15: Simulation results for the MI-GEE based on MNI and OIM methods ($K = 4$ - Skewed distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	109.3	0.298		107.3	0.289		111.7	0.325		106.4	0.303		117.8	0.338		108.7	0.303		117.0	0.036		111.1	0.070	
		β_t	108.3	0.032		116.9	0.032		99.8	0.033		114.9	0.035		107.7	0.033		117.0	0.036		117.0	0.036		111.1	0.070	
		β_{tx}	109.4	0.060		107.3	0.060		111.8	0.065		106.7	0.068		121.1	0.068		111.1	0.070		111.1	0.070		111.1	0.070	
3	300	β_x	102.9	0.103		101.2	0.100		108.0	0.106		102.9	0.098		110.9	0.120		102.2	0.106		110.9	0.120		102.2	0.106	
		β_t	91.6	0.011		100.6	0.011		90.4	0.010		107.2	0.011		96.9	0.011		113.0	0.011		96.9	0.011		113.0	0.011	
		β_{tx}	101.4	0.020		100.1	0.020		108.5	0.019		103.6	0.019		113.0	0.022		103.1	0.023		113.0	0.022		103.1	0.023	
3	500	β_x	102.5	0.060		100.9	0.058		105.6	0.063		100.7	0.058		110.0	0.072		101.5	0.061		110.0	0.072		101.5	0.061	
		β_t	91.9	0.006		101.8	0.006		86.2	0.006		103.7	0.006		96.2	0.006		112.3	0.007		96.2	0.006		112.3	0.007	
		β_{tx}	102.2	0.010		101.1	0.010		105.2	0.011		100.6	0.011		112.4	0.012		103.1	0.013		112.4	0.012		103.1	0.013	

Table 16: Simulation results for the MI-GEE based on MNI and OIM methods ($K = 5$ - Skewed distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	101.7	0.321		100.2	0.314		104.4	0.341		99.5	0.331		110.4	0.388		101.0	0.359							
		β_t	84.2	0.029		92.1	0.029		73.4	0.034		87.0	0.035		61.4	0.038		70.7	0.040							
		β_{tx}	101.3	0.060		99.9	0.060		105.6	0.067		100.0	0.069		115.0	0.080		103.2	0.082							
3	300	β_x	99.9	0.097		98.4	0.095		101.9	0.106		97.5	0.101		106.6	0.122		97.9	0.113							
		β_t	86.8	0.009		94.4	0.008		73.8	0.010		87.9	0.010		63.2	0.012		76.3	0.013							
		β_{tx}	100.9	0.019		99.4	0.019		102.5	0.021		98.0	0.021		110.2	0.025		99.1	0.026							
3	500	β_x	100.3	0.056		98.9	0.054		103.7	0.062		99.4	0.058		107.0	0.071		98.7	0.063							
		β_t	86.4	0.005		94.9	0.005		77.1	0.006		91.4	0.006		62.5	0.008		76.7	0.008							
		β_{tx}	100.3	0.011		99.1	0.011		104.7	0.012		100.2	0.012		109.5	0.014		99.3	0.014							

Table 17: Simulation results for the MI-GEE based on MNI and OIM methods (K = 7 - Skewed distribution)

		0%						10%						30%						50%						
T	N	Parm	MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM			MNI			FCS OIM		
			RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE		RB(%)	MSE	
3	100	β_x	102.5	0.285		101.5	0.277		105.8	0.307		100.6	0.281		109.8	0.337		99.5	0.289		109.8	0.337		99.5	0.289	
		β_t	81.6	0.028		91.2	0.028		69.2	0.030		88.9	0.029		52.6	0.032		76.7	0.032		52.6	0.032		76.7	0.032	
		β_{tx}	97.4	0.056		96.8	0.056		102.8	0.060		96.4	0.060		109.3	0.066		95.1	0.064		109.3	0.066		95.1	0.064	
3	300	β_x	102.7	0.093		101.4	0.091		105.8	0.102		101.7	0.094		108.4	0.114		100.5	0.099		108.4	0.114		100.5	0.099	
		β_t	92.1	0.008		100.9	0.008		78.2	0.009		102.5	0.009		57.4	0.011		88.0	0.010		57.4	0.011		88.0	0.010	
		β_{tx}	102.8	0.016		101.8	0.016		106.6	0.018		102.2	0.018		110.4	0.021		99.6	0.020		110.4	0.021		99.6	0.020	
3	500	β_x	102.8	0.058		101.6	0.057		106.5	0.064		102.4	0.058		109.6	0.074		102.1	0.063		109.6	0.074		102.1	0.063	
		β_t	88.9	0.005		97.8	0.005		76.3	0.006		99.9	0.005		56.5	0.008		88.7	0.006		56.5	0.008		88.7	0.006	
		β_{tx}	102.6	0.011		101.4	0.011		107.4	0.012		102.8	0.011		112.9	0.014		102.6	0.013		112.9	0.014		102.6	0.013	