Multiple imputation for ordinal longitudinal data with monotone missing data patterns

A.Y. Kombo, H. Mwambi & G. Molenberghs

Published online: 25 Apr 2016.

Submit your article to this journal ⎘

View related articles ⎘

View Crossmark data ⎘

Taylor & Francis
Taylor & Francis Group

Multiple imputation for ordinal longitudinal data with monotone missing data patterns

A.Y. Kombo[a,c], H. Mwambi[a] and G. Molenberghs[b]

[a]School of Mathematics, Statistics & Computer Science, UKZN, Pietermaritzburg, South Africa; [b]I-BioStat, Hasselt University and KU Leuven, Belgium; [c]Department of Mathematics and Informatics, Taita Taveta University College, Kenya

ABSTRACT

Missing data often complicate the analysis of scientific data. Multiple imputation is a general purpose technique for analysis of datasets with missing values. The approach is applicable to a variety of missing data patterns but often complicated by some restrictions like the type of variables to be imputed and the mechanism underlying the missing data. In this paper, the authors compare the performance of two multiple imputation methods, namely fully conditional specification and multivariate normal imputation in the presence of ordinal outcomes with monotone missing data patterns. Through a simulation study and an empirical example, the authors show that the two methods are indeed comparable meaning any of the two may be used when faced with scenarios, at least, as the ones presented here.

1. Introduction

Longitudinal studies are an important source of information in health sciences and other areas but often have the problem of missing data. Ordinal outcomes are increasingly becoming common in these studies. However, analysts are challenged if they need to impute missing values for such outcomes due to their hierarchical nature [11,12]. Missing values in longitudinal studies occur when not all of the planned measurements of a subject outcome vector are actually observed. This turns the statistical analysis into a missing data problem. For example, a subject may terminate early from a scheduled sequence of clinical visits for a number of reasons, both known and unknown. This type of missing pattern is termed dropout (monotone missing data pattern). Alternatively, a subject may miss a scheduled visit but appear at the next occasion. This is referred to as an arbitrarily (intermittent) missing data pattern. In this study, we focus on the former pattern of missingness. The reasons that lead to missingness are varied and it is always necessary to reflect on the nature of missingness and its impact on inferences. In [45], these reasons are classified into three categories. Data are said to be missing completely at random (MCAR) if the probability of missingness is independent of both the observed and unobserved measurements,

---

CONTACT A. Y. Kombo ✉ abdkombo@ttuc.ac.ke

missing at random (MAR) if, conditional on the observed data the probability of missing-ness is independent of the unobserved measurements and missing not at random (MNAR) for a violation of the above scenarios. Under the unrealistic MCAR, simple incomplete data methods such as last observation carried forward (LOCF), complete case analysis and available case analysis may be employed. However, even under the strong MCAR assumption it is not guaranteed that LOCF analysis is valid. In fact, analysts see it unscientific to use the ad hoc methods when broadly valid likelihood analyses can be easily implemented with standard software [9]. Generally speaking, the MAR assumption represents the most general condition under which valid inferences can be obtained without reference to the missing data mechanism, given inferences are likelihood based or Bayesian [9,27].

Recent advances in computational statistics have produced a new billow of flexible and formally justifiable procedures with sound statistical basis like multiple imputation (MI). MI, initially proposed by Rubin [46] and further detailed in [47,49], has become one of the most popular approaches in handling missing data. MI can be used not only with continuous variables but also with binary and categorical variables. It provides a way of accounting for uncertainty associated with imputations. This is a major strength against a number of existing single imputation methods. MI replaces each of the missing values with $m \geq 2$ plausible values generated under an appropriate imputation model to obtain m complete datasets. This replication captures the uncertainty about the missing data. The resulting m multiply imputed datasets are then analysed separately using an appropriate well-known standard method for complete data. The third stage is to combine the m analysis results into one for inferences, where the standard errors of estimates take account of the variation within and between the m imputations [47].

MI is a viable candidate for handling missing data in multivariate analysis. This is because it introduces appropriate random error into the imputation process and makes it possible to produce unbiased estimates of all parameters [2,47]. It can be used with any kind of data and any kind of analysis with/ without specialized software [2]. However, one key feature of MI is that, for correct and valid inferences, the imputation model should be correctly specified. It is agreed that the analysis and the imputation model should be congenial in the sense that the imputation model should be able to reproduce the major features of the analysis model [3,34,47]. In this paper, the imputation model includes the same variables that are in the analysis model. Regarding MI, it is also important to note that standard MI procedures assume that the data are MAR. While it is almost always impossible to test this assumption, including auxiliary variables in the imputation model that predict the missingness, together with variables that are correlated that will be included in the analysis model, can minimise bias. It also makes the MAR assumption more viable [14,50]. On the other hand, it is also possible to use MI procedures to impute data that are MNAR, but this requires making additional assumptions about the missingness mechanism.

This paper is concerned primarily with the comparison of two MI methods namely fully conditional specification (FCS) and multivariate normal imputation (MVNI) as applied to ordinal outcome variables with a monotone missing data pattern. Moreover, for the purpose of this paper, we focus on one ordinal outcome variable over time but the ideas presented here are applicable to other ordinal forms and data settings.

The paper is organised as follows. In Section 2, we give the key definitions and necessary notation. A description of the imputation methods is given in Section 3 followed by a simulation study and application in Section 4.

2. Definitions and notation

2.1. Missing data model

Suppose that for the ith subject in the study, a sequence of measurements Y_{ij} is expected to be measured at occasions $j = 1, \ldots, n_i$. Due to some reasons, some values of $Y_i = (Y_{i1}, \ldots, Y_{in_i})'$ are not observed. Then, Y_i can be partitioned into two subvectors such that $Y_{i,o}$ contains the observed measurements and $Y_{i,m}$ the unobserved measurements. Now, if we let Y to be the complete set of observations, then Y can be partitioned such that $Y = (Y_o, Y_m)$. We define a random vector $R_i = (R_{i1}, R_{i2}, \ldots, R_{in_i})$ compatible with the vector of observations Y_i such that $R_{ij} = 1$ if the outcome Y_{ij} is observed and 0 otherwise. Using [24], the joint distribution of the full data Y and the indicator vector variable R can be factorized as

$$f(Y, R \mid \theta, \psi) = f(Y \mid \theta) P(R \mid Y, \psi), \tag{1}$$

where ψ denotes a vector of parameters governing the missingness mechanism and θ denotes the measurement process model parameters. The conditional distribution of the missing data mechanism can be equivalently expressed as $f(R \mid Y_o, Y_m, \psi)$. Diggle and Kenward [22] propose modelling the probability of missingness at a particular measurement occasion as a linear function of the response values at previous occasions. For simplicity, we assume that this dropout depends only on the observed response just before the time it fails to be recorded and the unobserved response at the missing point. However, this model can be extended to include measured or observed covariates. If we denote by Y_{ij}, the response at measurement occasion j, the missing data model can be written as

$$\text{logit}[P_j(R_{ij} = 0 \mid y_{i1}, y_{i2}, \ldots, y_{i(j-1)}, y_{ij})] = \psi_0 + \psi_1 y_{i(j-1)} + \psi_2 y_{ij}, \tag{2}$$

where $P_j(R_{ij} = 0 | y_{i1}, y_{i2}, \ldots, y_{i(j-1)}, y_{ij})$ is the conditional probability of missingness at occasion j, given the history of responses, $y_{i1}, y_{i2}, \ldots, y_{i(j-1)}, y_{ij}$, the response subject to missingness, y_{ij} and ψ_0, ψ_1 and ψ_2 are the model parameters to be estimated. The model reduces to a MAR model if $\psi_2 = 0$. MCAR if $\psi_1 = \psi_2 = 0$. If $\psi_2 \neq 0$, then we cannot rule out MNAR but note that the test for $\psi_2 = 0$ versus $\psi_2 \neq 0$ (MAR versus MNAR) relies on untestable assumptions such as the distributional form [26,36,38,41]. In fact, Molenberghs et al. [35] show that a formal distinction between MAR and MNAR is not possible because for any MNAR model there exists a MAR counterpart that fits the data equally well.

2.2. Ordinal responses

There are cases where the outcome variable can be polytomous. While the typical logistic regression analysis models a binary response, logistic regression can also be applied to multilevel cases. If the response variable takes on values that have no inherent order (e.g. voting party A, B, C, or D), then the response is nominal. If it takes on intrinsic values like the levels of agreement (e.g. strongly agree to strongly disagree), then the response is ordinal. Then, for ordered categorical variables, the binary logistic regression extends to

polytomous logistic regression. A number of logistic regression models have been studied for ordinal response variables [1,6,15,31,33]. When there is need to consider several factors, special multivariate analysis for ordinal data is the natural alternative [17], although other methods, like mixed models may be used. However, ordinal logistic regression models have been most useful [5,33]. Several ordinal logistic regression models exist, namely the proportional odds model, partial proportional odds model (PPOM), continuous ratio model and the stereotype regression model. The most common among the ordinal logistic regression models is the proportional odds model [8]. The proportional odds model (a specific form of cumulative odds model) is a logit model that allows ordered data to be modelled by analysing it as a number of dichotomies. A binary logistic regression model compares one dichotomy (yes/no), whereas the proportional odds model compares a number of dichotomies by arranging the ordered categories into a series of binary comparisons. Here, the assumption is made that the effect of each explanatory variable is the same for each binary comparison (logit). This is the proportional odds assumption, also referred to as the parallel lines assumption (or equal slopes assumption). It leads to parsimony of the model, because it means that the effect of a predictor variable on the ordinal response is explained by one parameter. However, it may pose a restriction on the flexibility of the model, which may or may not be adequate for the data. Then before any model statistics are interpreted, it is important to test the assumption, a violation of which may lead to incorrect interpretation of results [5]. The assumption is commonly used with the cumulative logit link. On the other hand, mixed effects models have also been found very useful for longitudinal categorical (nominal or ordinal response) data. The main reason why random effects are used is to take account of correlated data due to clustering as a result of repeated measures from the same individual.

In medical and clinical research, it is not easy to get a continuous outcome for that kind of information you need. More often, the variable of interest has a natural ordering, say no disease, mild and severe. In this case using an ordinal outcome for the disease model may make sense other than 'no disease' and 'diseased', that is, collapsing the ordinal levels to binary ones. If this is done, one has to find an appropriate correlation structure of the dichotomized data, and then inflate the correlations intentionally in order to make them what they should have been. This means that one follows the ordinal-binary-Gaussian-ordinal-binary conversion scheme. This strategy, however, may not be applicable in every scenario [18]. The polytomous logistic regression model may be employed for the ordered categorical variable, but fails to make proper use of the information about the ordering. One way of taking advantage of the ordering is the use of 'cumulative odds', 'cumulative probabilities' and 'cumulative logits'.

Now, suppose that our data comprise a set of $i = 1, \ldots, N$ independent clusters (subjects in our longitudinal data context) where the ith subject consists of n_i observations. As before, let Y_{ij} denote the jth ($j = 1 \ldots, n_i$) response in subject i. This response may fall in any of $c = 1, \ldots, C$ distinct ordered categories for $C \geq 2$. Further, let x_{ij} denote a vector of predictor variables for the jth observation in the ith subject. Then, Y_{ij} will have a multinomial distribution with parameter vector $\boldsymbol{\pi}$. In this case, π_{jc} is the probability of the jth measurement falling into category c so that we have our cumulative probabilities given as

$$P(Y_{ij} \leq c) = P(Y_{ij} \leq c \,|\, x_{ij}) = \pi_{i1} + \cdots + \pi_{ic}. \tag{3}$$

Now using a logit link, we will have a cumulative logit model defined as

$$\text{logit}(P(Y_{ij} \leq c)) = \log\left[\frac{P(Y_{ij} \leq c)}{1 - P(Y_{ij} \leq c)}\right] = \alpha_c - x'_{ij}\boldsymbol{\beta}, \tag{4}$$

where $P(Y_{ij} \leq c)$ is the probability of being at or below category c, given a set of predictors. Here, $c = 1, \ldots, C-1$ for the C categories of the ordinal outcome, α_c gives the threshold parameters (intercept terms that depend on the categories). These parameters, however, are seldom of practical importance except for computing response probabilities. The regression parameters, $\boldsymbol{\beta}$, reflect the association between the predictor variables and the outcome variable. Notice that, while the regression coefficients do not vary (i.e. β has the same effect for each of the $C-1$ cumulative logits, implying that $x'_{ij}\boldsymbol{\beta}$ is independent of c), a different intercept exists for each level of the cumulative model. Given that the regression parameters ($\boldsymbol{\beta}$) are subtracted (model (4)), this means that a unit increase in the predictor variable will increase the log-odds of being in category greater than c. In other words, it means that the higher the value of $X'_{ij}\boldsymbol{\beta}$, the higher the probability of response falling in a category at the upper end of the response scale. But note that β itself can be estimated as negative which will give an increasing effect of the odds in categories less than or equal to c. The model describes the cumulative logits across $c-1$ response categories. One can transform the cumulative logits to obtain estimated cumulative odds and also the cumulative probabilities of being at or below category c.

3. Imputation methods

When the dataset has a monotone missingness pattern, variables with missing values are imputed sequentially with covariates obtained from their corresponding sets of preceding variables. To impute continuous variables, a regression method, a predictive mean matching method or a propensity score method may be used. A logistic regression method may be used for a binary or ordinal variable. Alternatively, a discriminant function for nominal or binary variables can be used. For real and simulated incomplete ordinal datasets, we contrast two multiple imputation procedures: the fully conditional specification (FCS) via chained equations [56,58] and the multivariate normal imputation (MVNI) [49]. These approaches are based on different theoretical assumptions and involve very different computational methods [28].

3.1. Multivariate normal imputation

Approaches to imputing multivariate data have been developed. For example, Rubin and Schafer [48] provided procedures for generating multivariate multiple imputation. This Bayesian simulation algorithm draws imputations from the posterior predictive distribution of the unobserved data given the observed data. The method assumes that the data are multivariate normally distributed and missing at random. Schafer [49] used this underlying approach and derived imputation algorithms for multivariate numerical, categorical and mixed data. The methodology describes the data by encompassing a multivariate model and derive a posterior distribution and then draw imputations from these by Gibbs sampling (here after referred to as data augmentation rather than Gibbs sampling). It uses the

Markov chain Monte Carlo (MCMC) approach to draw imputed values from the estimated multivariate normal distribution.

Given our ordinal response variable $Y \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, data augmentation [54] in Bayesian inference with missing data is based on iterating between an imputation step (I-step) and a Posterior step (P-step).

- *The imputation step* – With some estimated initial values for the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, the I-step simulates a value for missing data Y_m by randomly drawing it from the conditional predictive distribution of Y_m, that is, from a current estimate (rth iteration) $\theta^{(r)}$, of the parameter, a value Y_m^{r+1} of the missing data is drawn from the conditional distribution of Y_m given Y_o:

$$Y_m^{(r+1)} \sim P(Y_m \mid Y_o, \theta^r), \quad \boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{5}$$

- *The posterior step* – This step draws a value of the parameter $\boldsymbol{\theta}$ from a complete-data posterior distribution:

$$\theta^{(r+1)} \sim P(\theta \mid Y_o, Y_m^{(r+1)}). \tag{6}$$

The updated estimates are then used in the imputation step.

Iterating Equations (5) and (6) from initial value $\theta^{(0)}$ will yield a stochastic sequence $\{(\theta^{(r)}, Y_m^{(r)}); r = 1, 2, \ldots\}$. The two steps are iterated sufficiently long until the distribution of the estimates becomes stationary [49]. Each step depends on the previous one, meaning that there is dependency across the steps. This approach is theoretically sound but based on distributional assumptions that may not always be realistic (e.g. assuming normality for binary, ordinal variables). For categorical variables, the MVNI method draws imputations under the MVN model and so we need to round off the imputations to the nearest integer to accommodate the categorical nature of the data. Allison [4], however, cautions about rounding (he cites the binary case) because the rounded imputed values may lead to biased parameter estimates. Nonetheless, Schafer [49] still argues that inference from MVNI may be reasonable even if multivariate normality does not hold, for example, in the cases of binary and categorical variables. We refer the reader to [49] for a detailed account of this procedure.

3.2. Fully conditional specification

An alternative option, applicable to multivariate data, is the fully conditional specification (FCS) approach. FCS is a flexible method that specifies the multivariate model by a series of conditional models for each of the incomplete variables. Unlike MVNI, it does not necessarily rely on the multivariate normality assumption and thus univariate regression models can be appropriately tailored to be used for ordered logistic regression for ordinal variables. Using a Bayesian approach, imputations are done stepwise starting with the variable with the least amount of missing values and progressing like that until the variable with the most missing data is finally handled. It involves two phases in each imputation: the filled-in stage and the imputation stage. During every stage, draws are randomly done from both the posterior distribution of the parameters and posterior distribution of the missing values. At

the filled-in stage, the missing values are filled in sequentially over the variables, one after the other with preceding variables serving as covariates. The filled-in values are then used as starting values for the imputation stage. At the imputation stage, the filled-in values are replaced with imputed values for each variable sequentially at each iteration.

Let the ordinal response variable Y be characterized by a vector of unknown parameters $\theta = (\mu, \Sigma)$; μ is a mean vector while Σ is a covariance matrix. As before, $Y = (Y_o, Y_m)$. Following [59] and also in [10], multiple imputation via FCS proceeds as follows:

- calculate the posterior distribution of θ given the observed data, that is, $P(\theta \mid y_o)$;
- draw a value θ^* from $P(\theta \mid y_o)$;
- draw a value y^* from the conditional posterior distribution of y_m given $\theta = \theta^*$:

$$y^* \sim P(y_m \mid y_o, \theta = \theta^*). \tag{7}$$

Repeat the second and third steps depending on the number of imputations. The steps are repeated long enough for the results to reliably simulate an approximately independent draw of the missing values for an imputed dataset.

3.3. Software considerations

When we assume MAR, valid inferences can be obtained through likelihood-based analysis without modelling the dropout process. Consequently, the generalized linear mixed model – as the analysis model – is used. This approach may be implemented by using SAS procedures NLMIXED and GLIMMIX. If we need to impute missing values, both the description of missing data patterns and multiple imputation is performed using the procedure PROC MI. It may be used for all types of variables. The procedure offers several methods for imputation depending on whether the variable is continuous or categorical. Here, we are interested in comparing MVNI and FCS as implemented in PROC MI. For MVNI, it uses the Markov Chain Monte Carlo (MCMC) approach to draw imputed values from the estimated multivariate normal distribution. To use it, the user calls it by specifying the mcmc statement in the MI procedure. To run FCS, the fcs statement is specified in PROC MI. In PROC MI, the imputation model to be used and the number of imputed datasets to be created are specified. After imputation, statistical procedures run the analytic model of interest separately for each imputation using _Imputation_ as a BY variable, and the results are stored in an output file. Finally, a procedure call, PROC MIANALYZE, combines the estimates obtained from the analyses for multiply imputed data to produce valid statistical inferences. However, for some complete data analyses, like those for categorical data, additional manipulations are needed before PROC MIANALYZE is used [40]. This is because Rubin rules [47] for combining results assume that the statistics estimated are normally distributed. Such estimates, like regression coefficients and means, are approximately normally distributed, while others like the odds ratios, correlation coefficients and relative risks are nonnormal. If interest is on the latter group of estimates, they can first be normalized before applying Rubin's combination rules to the transformed estimates. In [57], some transformations to various types of estimated statistics are suggested.

By default, the SAS procedure LOGISTIC fits the proportional odds model combined with the cumulative logit link. When the assumption of the common slopes is valid for some variables but not for others, PROC GENMOD may be used to fit the

PPOM. Alternatively, PROC LOGISTIC may also be used but with a specification of the UNEQUALSLOPES option in the model. PROC CATMOD can be used in case of a nonproportional odds model.

4. Simulation study

4.1. Data generation, simulation designs and analysis of the simulated data

We conducted a simulation study to examine the performance of FCS and MVNI. The datasets were generated using a scenario that mimics common longitudinal studies. The simulated datasets are based on an ordinal outcome with C categories which are generated at four study occasions, $j = 1, \ldots, 4$. The setting was repeated for three different settings where $C = 3, 4, 5$. For each of the different scenarios, we simulated 1000 datasets based on a generalised linear mixed model scheme of the form (8) for sample sizes $N = 100, 250, 500$. Consequently, longitudinal ordinal variables were generated following a model with a linear predictor:

$$\text{logit}[P(Y_{ij}^* \leq c)] = \alpha_c + x'\beta + b_i, \quad b_i \sim N(0, d). \tag{8}$$

An ordinal regression model was motivated by assuming an underlying latent variable (y^*) which is related to the actual response through the 'threshold concept'. The response is defined based on some underlying unobserved continuous endpoint that follows a linear regression model incorporating random effects and a prespecified set of cut-off values (threshold values) α_c. The data were generated by assuming a vector of predictor variables $x' = (x_1, x_2, x_3, x_4)$, which is a combination of both continuous and binary variables. Here, x_1 and x_3 are binary group effects (i.e. $x = 0,1$) representing a treatment group indicator and gender respectively, x_2 is a continuous variable representing exposure period and x_4 is a four-point assessment time. For the simulations we used the parameters, $\beta_1 = 0.9, \beta_2 = 0.2, \beta_3 = 0.5$ and $\beta_4 = 0.8$. For simplicity of the simulations in this paper, we did not assume any interaction of terms. In this case, our simulation model is explicitly written as

$$\text{logit}[P(Y_{ij}^* \leq c)] = \alpha_c + 0.9x_1 + 0.2x_2 + 0.5x_3 + 0.8x_4 + b_i, \quad b_i \sim N(0, 1.8^2). \tag{9}$$

By inverting the logit link function, it leads to the conditional ordinal logistic regression model, noting that Equation (8) can be equivalently written as

$$P(Y_{ij}^* \leq c) = \frac{\exp(\alpha_c + x'\beta + b_i)}{1 + \exp(\alpha_c + x'\beta + b_i)}. \tag{10}$$

Let $\phi_{ijc} = P(Y_{ij}^* \leq c)$, we obtain the ordinal response Y_{ij} (e.g. for $C = 4$) by setting an observation rule defined as

$$Y = \begin{cases} 1 & \text{if } \phi_{ij} \leq \tau_1, \\ 2 & \text{if } \tau_1 < \phi_{ij} \leq \tau_2, \\ 3 & \text{if } \tau_2 < \phi_{ij} \leq \tau_3, \\ 4 & \text{if } \phi_{ij} > \tau_3. \end{cases} \tag{11}$$

First from the full datasets without imposing any missing values, parameters and standard errors were estimated by a likelihood based approach. Each estimate is an average of

1000 estimates from the different simulated datasets. Then, we assumed a rather simple MAR model of missingness, where subjects whose outcome was greater than some cut-off probability would miss at post baseline time points 3 and 4, that is, let $drp = y_{ij} - y_{ij-1}$, $j = 2, 3, 4$, yielding values between -2 and 2; -3 and 3; and -4 and 4 for the different choices of the categories of the ordinal outcome, that is, for $C = 3, 4$ and 5 categories respectively. Then, we normalized these values by defining $ndrp = (drp + (C - 1))/2C$ in order to confine them to the range $[0, 1]$. Finally, if $ndrp > \gamma + 0.6u$ (where $u \sim [0, 1]$ is a uniformly distributed random number), then $y_{i(j+1)}$ misses. We held (for the $C = 3,4,5$ categories, respectively) $\gamma = 0.4$ so as to ensure that about 30% of the response data were missing. The probability of a value dropping depended merely on the immediate history.

Then, the missing entries were imputed using FCS and MVNI as carried out in PROC MI. We used the expectation-maximization (EM) algorithm [20] to obtain the starting values for our imputations. MVNI was performed using the SAS PROC MI with a specification of the MCMC statement. The ordinal values were imputed on the continuous scale and rounded off to the required categories. Maximum and minimum values were specified based on the scale of the response options of the dataset. These specifications were necessary so as to ensure that imputations were not created outside the range of the response values. FCS was carried out using fcs statement in PROC MI. The ordinal response was imputed using the ordinal logistic regression model as incorporated in the FCS procedure. For all cases in the study, default values for MCMC and FCS specifications were used in the simulations. We realized that the algorithms still converged to the correct posterior distributions and were confident that the imputed values in the different datasets were statistically independent. All the other predictor variables were used to ensure that our imputation model was rich enough to try and satisfy the congeniality requirement under the MAR assumption. For simplicity, throughout the analyses in this paper the categorical time was treated as continuous.

For comparison of methods, a larger number of imputations are necessary [61]. We performed $m = 20$ imputations. This relatively high value was chosen to account for the relatively large fraction of missing data and to limit the loss of power for testing any associations of interest. Nonetheless, researchers argue that m can be set to $3 \leq m \leq 5$ and still get sufficient accuracy. However, Schafer [49] cautions that pegging on this range might be risky. On the other hand, Molenberghs and Verbeke [37] showed that efficiency increments diminish rapidly after the first $m = 2$ imputations for a small fraction of missing information and after the first $m = 5$ imputations for larger fractions of missing information. However, a rule of thumb for choosing m is suggested (see [60]). They suggest that m should be at least equal to the percentage of incomplete cases. Nevertheless, we caution the reader that still discretion is necessary, based on the problem at hand.

To compare the performance, we used bias and mean squared error (MSE) of the parameter estimates. We defined bias as the absolute difference between the average parameter estimate from the analysis procedures (based on the 1000 data replications) and the true value (i.e. Bias $= |\bar{\hat{\beta}} - \beta|$).

4.2. Simulation results

Results of the simulation study (based on 1000 simulated datasets and 20 imputations) are presented. We present three tables, where Table 1 represents results when the ordinal

Table 1. Standard errors (Std Err), Bias and mean squared error (MSE) estimates from fully conditional specification (FCS) and multivariate normal imputation methods (MVNI).

Sample	Par	Std Err				Bias				MSE			
		FDA	DL	FCS	MVNI	FDA	DL	FCS	MVNI	FDA	DL	FCS	MVNI
$N = 100$	β_1	0.0114	**0.0130**	0.0111	0.0106	0.0614	0.0154	0.0144	**0.0965**	0.0039	0.0004	0.0003	**0.0093**
	β_2	0.0007	**0.0008**	0.0007	0.0007	0.0108	0.0042	0.0043	**0.0212**	0.0001	0.0000	0.0000	**0.0004**
	β_3	0.0128	**0.0158**	0.0125	0.0120	0.0269	0.0100	0.0074	**0.0556**	0.0009	0.0003	0.0002	**0.0032**
	β_4	0.0030	0.0035	**0.0040**	**0.0040**	0.0559	0.1820	0.1826	**0.2354**	0.0031	0.0331	0.0334	0.0554
$N = 250$	β_1	0.0076	**0.0091**	0.0077	0.0075	0.0540	0.0191	0.0193	**0.1004**	0.0030	0.0004	0.0004	**0.0101**
	β_2	0.0004	**0.0006**	0.0005	0.0004	0.0113	0.0048	0.0049	**0.0217**	0.0001	0.0000	0.0000	**0.0005**
	β_3	0.0082	**0.0094**	0.0077	0.0074	0.0356	0.0157	0.0167	**0.0620**	0.0013	0.0003	0.0003	**0.0039**
	β_4	0.0018	0.0022	0.0024	**0.0028**	0.0547	0.1821	0.1824	**0.2350**	0.0030	0.0332	0.0333	0.0552
$N = 500$	β_1	0.0056	**0.0067**	0.0053	0.0059	0.0611	0.0274	0.0272	**0.1086**	0.0038	0.0008	0.0008	**0.0118**
	β_2	0.0003	**0.0004**	0.0003	0.0003	0.0111	0.0048	0.0049	**0.0218**	0.0001	0.0000	0.0000	**0.0005**
	β_3	0.0052	**0.0065**	0.0053	0.0052	0.0440	0.0246	0.0246	**0.0693**	0.0020	0.0006	0.0006	**0.0048**
	β_4	0.0013	0.0016	**0.0018**	0.0017	0.0554	0.1843	0.1844	**0.2369**	0.0031	0.0340	0.0340	0.0561

Notes: Also estimates from full data analysis (FDA) and direct likelihood (DL) method. Missing values, approximately (30%) on the response variable; MAR mechanism. A case where ordinal variable has $C = 3$ levels.

Table 2. Standard errors (Std Err), Bias and mean squared error (MSE) estimates from fully conditional specification (FCS) and multivariate normal imputation methods (MVNI).

Sample	Par	Std Err				Bias				MSE			
		FDA	DL	FCS	MVNI	FDA	DL	FCS	MVNI	FDA	DL	FCS	MVNI
$N = 100$	β_1	0.0100	**0.0118**	0.0100	0.0099	**0.2382**	0.1598	0.1595	0.1933	**0.0568**	0.0257	0.0255	0.0375
	β_2	0.0006	**0.0007**	0.0006	0.0006	**0.0517**	0.0314	0.0315	0.0358	**0.0027**	0.0010	0.0010	0.0013
	β_3	0.0102	**0.0123**	0.0106	0.0097	**0.1214**	0.0879	0.0894	0.1082	**0.0148**	0.0079	0.0081	0.0118
	β_4	0.0023	0.0033	**0.0036**	**0.0036**	0.2141	0.3929	0.3930	**0.4007**	0.0458	0.1544	0.1545	**0.1606**
$N = 250$	β_1	0.0063	**0.0079**	0.0065	0.0064	**0.2329**	0.1556	0.1563	0.1914	**0.0543**	0.0243	0.0245	0.0367
	β_2	0.0003	**0.0004**	**0.0004**	0.0003	**0.0520**	0.0316	0.0316	0.0359	**0.0027**	0.0010	0.0010	0.0013
	β_3	0.0066	**0.0082**	0.0066	0.0066	**0.1290**	0.0868	0.0871	0.1082	**0.0167**	0.0076	0.0076	0.0118
	β_4	0.0015	0.0020	0.0021	**0.0022**	0.2123	0.3884	0.3885	**0.3972**	0.0451	0.1509	0.1509	**0.1578**
$N = 500$	β_1	0.0044	**0.0056**	0.0047	0.0045	**0.2384**	0.1611	0.1614	0.1958	**0.0569**	0.0260	0.0261	0.0384
	β_2	0.0002	**0.0003**	0.0003	0.0002	**0.0521**	0.0316	0.0317	0.0360	**0.0027**	0.0010	0.0010	0.0012
	β_3	0.0042	**0.0054**	0.0047	0.0045	**0.1385**	0.0952	0.0951	0.1164	**0.0192**	0.0091	0.0091	0.0136
	β_4	0.0010	0.0015	**0.0017**	0.0014	0.2131	0.3910	0.3913	**0.3989**	0.0454	0.1529	0.1531	**0.1591**

Notes: Also estimates from full data analysis (FDA) and direct likelihood (DL) method. Missing values, approximately (30%) on the response variable; MAR mechanism. A case where ordinal variable has $C = 4$ levels.

outcome variable has three categories/levels, Table 2, the variable has four levels and Table 3 when the variable has five levels. The results are presented for MVNI, FCS, direct likelihood (DL) and full data analysis (FDA). In this paper, full data refer to the simulated dataset that has no missing values. Although the original idea of the paper was to contrast the performance MVNI and FCS, DL is presented as an additional approach because of its known ability to handle incomplete data. Rather than imputing missing measurements, Mallinckrodt *et al.* [32] suggested the use of a direct likelihood approach to deal with incomplete correlated data under the ignorable assumption. Here, the observed cases are analysed without any analyst's adjustments, that is, without imputation nor deletion, by the use of models that provide a framework where clustered data can be analysed by including both fixed and random effects in the model (in case of GLMMs for non-Gaussian data) [25]. The authors in [25] further showed that DL analysis of incomplete datasets produced unbiased

Table 3. Standard errors (Std Err), Bias and mean squared error (MSE) estimates from fully conditional specification (FCS) and multivariate normal imputation methods (MVNI).

Sample	Par	Std Err				Bias				MSE			
		FDA	DL	FCS	MVNI	FDA	DL	FCS	MVNI	FDA	DL	FCS	MVNI
$N = 100$	β_1	0.0103	**0.0114**	0.0093	0.0096	0.2500	0.2295	0.2294	**0.2565**	0.0626	0.0528	0.0527	**0.0659**
	β_2	0.0005	**0.0006**	0.0005	0.0005	**0.0546**	0.0499	0.0499	0.0514	**0.0030**	0.0025	0.0025	0.0026
	β_3	0.0106	**0.0115**	0.0097	0.0094	0.1354	0.1253	0.1247	**0.1415**	0.0184	0.0158	0.0156	**0.0201**
	β_4	0.0024	0.0030	**0.0034**	0.0033	0.2228	0.2908	0.2904	**0.3221**	0.0496	0.0846	0.0843	**0.1038**
$N = 250$	β_1	0.0061	**0.0069**	0.0059	0.0056	0.2433	0.2279	0.2304	**0.2562**	0.0592	0.0520	0.0531	**0.0657**
	β_2	0.0003	0.0003	0.0003	0.0003	**0.0549**	0.0502	0.0503	0.0518	**0.0030**	0.0025	0.0025	0.0027
	β_3	0.0068	**0.0076**	0.0060	0.0064	**0.1409**	0.1255	0.1255	0.1408	**0.0199**	0.0158	0.0158	**0.0199**
	β_4	0.0013	0.0018	0.0018	**0.0019**	0.2230	0.2903	0.2904	**0.3228**	0.0497	0.0843	0.0843	**0.1042**
$N = 500$	β_1	0.0044	**0.0052**	0.0043	0.0041	0.2497	0.2320	0.2317	**0.2597**	0.0624	0.0539	0.0537	**0.0675**
	β_2	0.0002	**0.0003**	0.0002	0.0002	**0.0549**	0.0503	0.0503	0.0519	**0.0030**	0.0025	0.0025	0.0027
	β_3	0.0041	**0.0049**	0.0039	0.0041	0.1465	0.1328	0.1323	**0.1482**	0.0215	0.0177	0.0175	**0.0220**
	β_4	0.0010	0.0013	**0.0014**	**0.0014**	0.2244	0.2918	0.2922	**0.3241**	0.0504	0.0851	0.0854	**0.1050**

Note: Also estimates from full data analysis (FDA) and direct likelihood (DL) method. Missing values, approximately (30%) on the response variable; MAR mechanism. A case where ordinal variable has $C = 5$ levels.

parameter estimates that were comparable to those from a full data analysis. These arguments were echoed by Molenberghs and Verbeke [37], who also pointed out cases where MI is justified.

For clarity, results are presented here for regression coefficients only and not the intercepts. In all tables, larger values depicting worst cases are in bold.

In Table 1, considering bias, we notice that the largest values are obtained for MVNI. These are followed by FDA in all cases except β_4 where FCS produces larger values than FDA. The trend is the same for all sample sizes. The FCS and DL values are very close to each other with one case ($\beta_3, N = 500$) where they are the same. Looking at MSE, we observe a similar situation as for bias, that is, bigger values for MVNI followed by FDA except β_4 where FCS produces larger values than FDA. Comparing DL and FCS, we see equal values for all cases save for $\beta_1, \beta_3, \beta_4$ for $N = 100$, and β_4 in $N = 250$. However, these values are very close such that in 3-decimal places, they are equal. Looking at standard errors, largest values are observed for DL consistently except β_4. MVNI produces the smallest values in all the other cases except β_4, for $N = 100,250$.

Now shifting focus to Table 2, the scenario we observed in Table 1 changes. We notice that largest bias are recorded for FDA for all β's except β_4 where MVNI gives the largest bias. Exactly, the same trend is produced under MSE. Looking at standard errors, here the same scenario as in Table 1 is reproduced. Again, DL and FCS produce the same or very close values.

In Table 3, the previous trends observed for standard errors are replicated here. For bias and MSE, the trends change slightly. Now, the largest biases are recorded for MVNI in all cases except β_2 for all sample sizes, and β_3 for $N = 250$. The same set-up is produced under MSE. Like before very close or equal values are observed for FCS and DL.

In terms of bias MVNI seems to be more biased than FCS. If one is interested in smaller standard errors, then MVNI has mostly smaller values than FCS or at times they are equal. Generally, FCS may seem slightly better than MVNI, but both methods seem to perform equally well. DL is another favourable alternative in case one is not well conversant with the imputation methods. Faster and easily implemented in standard statistical software.

4.3. Example: arthritis data

4.3.1. Data

The dataset used is from a homoeopathic clinic in Dublin, made available in [39]. The data are on 60 patients (12 males and 48 females) between the ages of 18 and 88 who were under treatment for arthritis. The patients were followed up for a month (in 12 visits) and their pain scores assessed. Only two patients had all the scores for the 12 visits. The score was graded from 1 to 6, with high indicating worse. Only those with a baseline score greater than 3 and a minimum of six visits are reported. About 36% of the pain score data were missing. Of the 60 patients, 27 had RA-type arthritis where 5 were males and 22 were females, while 33 had type OA. Seven of these were males. Some descriptive statistics of the dataset are summarized in Table 4 and Figure 1.

Looking at Figure 1, it is apparent that many patients missed their visits towards the end of the follow up. After the sixth visit, the missing data were more than 30% on every visit.

Table 4. Descriptive statistics of the incomplete arthritis data.

Arthritis data Variable	Description	Range	% miss	Mean	Mode	Std Dev.
Baseline variables						
Sex	1 = Male, 0 = Female	1/0	0			
Age	Age of the patient	18–88	0	59.5	57	12.6
Time	Number of patient visits	1–12	0			
Type	Arthritis type[a] (RA = 1, OA = 0)	1/0	0			
Years	Number of years with symptom	0–57	0	10.7	1	12.2
Response variable						
pain_scores	Scores on the arthritis pain	1–6	35.56%		4	

Note: Data missing on the dependent variable.[a] Arthritis type (RA, rheumathoidarthritis, OA = ostheo − arthritis). StdDev., standarddeviation.
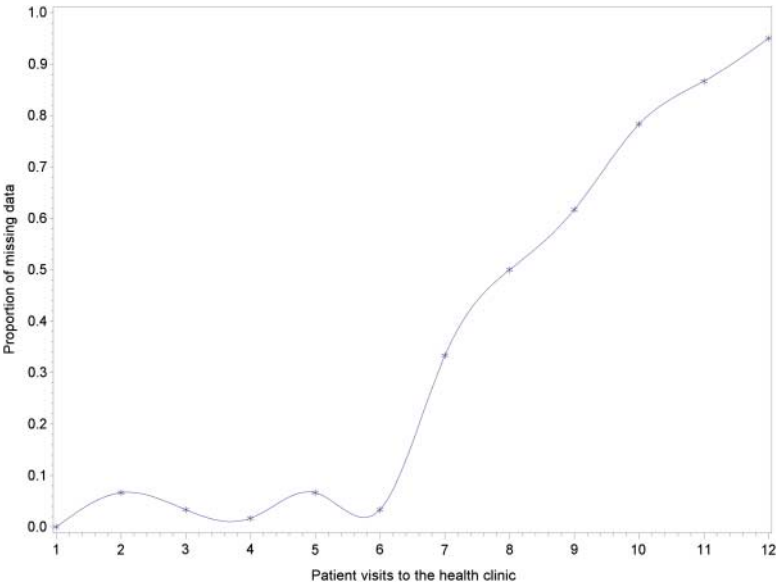

Figure 1. The proportion of missing data per scheduled visit to the clinic.

4.3.2. The proportional odds assumption

Before the model statistics can be interpreted, it is very important to test the proportional odds assumption. The assumption was examined using the Brant test in STATA. A non-significant omnibus test provides informal evidence that the assumption is not violated. Table 5 gives part of the assumption results. The assumption was upheld for age, type and years. The same cannot be said for sex and time.

A model of interest for the study was the main effects model. Only the dependent variable had missing values. At first, the data were analysed without any alterations or attempts to impute the missing values. This was under the direct likelihood (DL) approach. We chose the DL parameter estimates as reference for the real application dataset against which we can check the relative performance of MVNI versus FCS when considering MI. Because direct likelihood is valid under the same properties as multiple imputation, we expect the two approaches to produce similar parameter estimates or somehow close to each other. After the direct likelihood analysis, we conducted the multiple imputations under FCS and MVNI where upon imputation, a similar marginal model as the direct likelihood analysis was fitted in the analysis task. Finally, the SAS procedure MIANALYZE was employed to pool the results from multiple datasets.

4.3.3. Results

Table 6 shows the parameter estimates, standard errors and 95% confidence limits of fixed effect estimates by the imputation methods and direct likelihood analysis. These analysis results showed similar trends to those from the simulated data for most cases. The results indicate that the parameter estimates by MVNI were comparable to those of direct likelihood in more cases than FCS. In three cases, MVNI values were closer to those from the direct likelihood method compared to two FCS cases. Moreover, MVNI resulted in smaller standard errors than the FCS method for age, time and type. Equal values are observed for

Table 5. Brant test of proportional odds assumption.

Variable	chi2	$p >$ chi2	df
All	73.87	0.000	20
Sex	34.88	0.000	4
Age	7.59	0.108	4
Time	30.55	0.000	4
Type	8.35	0.079	4
Years	6.00	0.199	4.

Table 6. Parameter estimates, standard errors (StdErr) and confidence limits (C.L.) obtained from the arthritis data under the methods of direct likelihood (DL), fully conditional specification (FCS) and multivariate normal imputation (MVNI).

	DL			MVNI			FCS		
Param	Est	StdErr	95% C. L.	Est	StdErr	95% C. L.	Est	StdErr	95% C. L.
Sex	0.2130	0.1543	(−0.0895, 0.5154)	0.2192	0.2273	(−0.2282, 0.6667)	0.1991	0.2212	(−0.2361, 0.6342)
Age	0.0262	0.0062	(0.0140, 0.0383)	0.0260	0.0070	(0.0127, 0.0393)	0.0255	0.0072	(0.0114, 0.0396)
Time	−0.2218	0.0610	(−0.3414, −0.1023)	−0.2004	0.0284	(−0.2565, −0.1444)	−0.2212	0.0304	(−0.2812, −0.1612)
Type	0.9868	0.1275	(0.7369, 1.2366)	0.9620	0.1677	(0.6326, 1.2914)	0.9799	0.1770	(0.6319, 1.3280)
Years	0.0107	0.0045	(0.0018, 0.0195)	0.0106	0.0070	(−0.0032, 0.0243)	0.0099	0.0070	(−0.0038, 0.0236)

Note: Missing values about (36%) on the response variable.

years. MVNI gives a larger standard error than FCS for sex. This may be attributed to the fact that both sex and years were highly insignificant predictors by both MVNI and FCS, as is evidenced in the confidence limits. Both methods seem to perform fairly well in general. Looking at the direct likelihood method, it gives smaller standard errors than the imputation methods for all parameters except time. It is equally a favourable alternative method when faced with incomplete ordinal data and may be used whenever one is not sure about what imputation method to use or not having necessary know how on imputation methods.

5. Discussion

The idea behind MI is to draw valid and efficient inferences by fitting analysis models to multiply imputed data. We ensured that the imputed values bear the structure of the data, and uncertainty about the structure and included any knowledge about the process that led to the missing data [56]. The method of choice to create the imputed datasets depends on the missing data pattern. For monotone missing patterns, a parametric regression method that assumes multivariate normality or a nonparametric method that employs propensity scores may be used [37]. Alternatively, one may generate imputations by performing a series of univariate regressions, rather than just a single large model (making it somewhat easier to estimate), and without assuming normality of the variables.

When faced with a discrete variable (e.g. ordinal), an appealing approach at first sight may be to treat ordinal variables as continuous for the purpose of imputation, and then round the imputed data values to the nearest valid discrete value before continuing to fit the substantive model [11]. However, researchers caution the analyst from analysing ordinal outcome as a continuous or dichotomized variable for a number of reasons. First, comparing an ordinal to a continuous outcome or dichotomizing it to run a binary logistic regression may lead to efficiency loss due to information loss, reduced statistical power and decreased generality of the analytic conclusions [23]. Logically, continuous models can yield predicted values outside the range of the ordinal variable and finally, a continuous model may produce correlated residuals and regressors when used for ordinal outcomes and does not account for the ceiling and floor effects of the ordinal outcome. This may lead to biased estimates of the regression coefficients [7]. This issue has created a lot of debate among researchers. Schafer [49] argues that methods assuming multivariate normality may be used in cases where the normality assumption does not hold. Furthermore, these methods have also been successfully used by the authors in [13,19,53]. This is therefore still an active area of further research. However, apart from imputing the ordinal variable directly as a continuous variable, another option is to use a set of indicators. The values are imputed as continuous, and then assign imputed values into categories based on the mean indicators imputed in a separate round of imputation. In [29], this strategy of comparing methods for imputing ordinal data using methods that assume multivariate normality is discussed.

More often analysts are faced with datasets with both dropouts and nonmonotone missingness, like the arthritis data where the amount of dropout was considerable, while that of nonmonotone missingness is much smaller. It is heedful to include all in the analyses as noted by Molenberghs and Verbeke [37]. One can undisputedly opt for direct likelihood analysis or standard generalized estimating equation (GEE; [21,30,37]). Weighted generalized estimating equation (WGEE; [43]) is possible but one has to find appropriate weights.

Alternatively, one may make the missing patterns monotone by multiple imputation and go ahead to do the WGEE.

The primary goal for this study was to investigate the performance of MVNI and FCS as MI methods. These two approaches follow different theoretical assumptions and thus involve different computational methods. Each of the methods comes with its own specifications. MVNI is appealing because of its ease of specification of the imputation model. Conversely, FCS requires an added effort in model specification, and separate regression models must be fitted for each variable in the imputation model [56]. But in our problem these conditional regressions were automatically specified because of our small number of variables and only one variable had missing values. On the other hand, an added advantage of FCS again is the natural handling of ordinal variables. For MVNI, we had to handle the ordinal variables under a continuous scale in order to take advantage of the well-established imputation procedures for Gaussian outcomes, and then rounded to the required categories post-estimation. Basically, this assumption has been the major stumbling block in the working of MVNI and a number of researchers have reported FCS being better than MVNI, for example [56, 62]. In this study, we did not find a strong reason to support this. Specifically speaking, the MVNI approach is equally appropriate as is FCS when faced with missingness in ordinal variables, at least of the type presented. Similarly, Lee and Carlin [28] are in support of the findings. However, without doubt, further comparisons on these two methods, where more settings will be considered is incumbent.

In this paper, we focussed on MAR mechanisms for monotone missing data patterns. The methods of FCS and MVNI can be extended to non-monotone missing data patterns [55]. Although the authors doubt the suitability of the MAR assumption for non-monotone missing data, Robins and Gill [42] present a new strategy of ignorable non-monotone missing data models, called the randomised monotone missingness (RMM), which is a subset of MAR. They argue that the RMM is the only plausible non-monotone MAR mechanism that is not MCAR, but they caution the user not to analyse non-monotone missing data assuming that the missingness is ignorable if a statistical test has rejected the hypothesis that the missingness process can be represented as RMM. We recommend interested readers to [42] for further details on RMM and [16] who reiterate the RMM idea and extend it to a Markov randomised monotone missingness (MRMM). MRMM is a specific subset of RMM. The authors present a clear theoretical framework and applicability in non-monotone missingness patterns. We therefore state that the methods employed in our paper can further be extended to non-monotone cases. These methods are valid under MAR. When faced with non-monotone missingness, one may take the [16,42] routes as one of the options that exist in the literature. If under any circumstances, it happens that the MAR is not a sensible assumption for non-monotone missing cases, as an outset, sensitivity analyses are advised. However, a shift from MAR to possibly MNAR is not a worry, because as pointed out by Molenberghs *et al.* [35] the price to pay is minimal as no formal distinction exists between MAR and MNAR. This is because for any MNAR model there exists an MAR counterpart that fits the data very well.

In this paper, missingness was only on the outcome variable. This does not limit the applicability of FCS and MVNI to that case only. The methods can be extended to situations where data are missing for outcomes and covariates. A lot of work has been done on this. In the papers [44,58], MICE alias FCS was used to fill missing values in incomplete

covariates. The assumption of multivariate normality has been used to impute in covariates and responses. We cite [49,51,52] among many works in the literature.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] A. Agresti, *Tutorial on modeling ordered categorical response data*, Psychol. Bull. 105 (1989), pp. 290–301.

[2] P.D. Allison, *Multiple imputation for missing data: A cautionary tale*, Sociol. Methods Res. 28 (2000), pp. 301–309.

[3] P.D. Allison, *Missing Data*, SAGE, Thousand Oaks, CA, 2001.

[4] P.D. Allison, *Imputation of categorical variables with PROC MI*, in *SAS Users Group International, 30th Meeting (SUGI 30)*, 2005, April.

[5] C.V. Ananth and D.G. Kleinbaum, *Regression models for ordinal responses: A review of methods and applications*, Int. J. Epidemiol. 26 (1997), pp. 1323–1333.

[6] B.G. Armstrong and M. Sloan, *Ordinal regression models for epidemiologic data*, Am. J. Epidemiol. 129 (1989), pp. 191–204.

[7] D.J. Bauer and S.K. Sterba, *Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation*, Psychol. Methods 16 (2011), pp. 373–390.

[8] R. Bender and U. Grouven, *Using binary logistic regression models for ordinal data with non-proportional odds*, J. Clin. Epidemiol. 51 (1998), pp. 809–816.

[9] C. Beunckens, G. Molenberghs, and M.G. Kenward, *Direct likelihood analysis versus simple forms of imputation for missing data in randomized clinical trials*, Clin. Trials 2 (2005), pp. 379–386.

[10] S. Buuren and K. Groothuis-Oudshoorn, *MICE: Multivariate imputation by chained equations in R*, J. Stat. Softw. 45 (2011).

[11] J.R. Carpenter and M.G. Kenward, *Multiple Imputation and its Application*, Wiley, Chichester, 2013.

[12] L. Chen, M. Toma-Drane, R.F. Valois, and J.W. Drane, *Multiple imputation for missing ordinal data*, J. Mod. Appl. Stat. Methods 4 (2005), p. 26.

[13] K.-H. Choi, C. Hoff, S.E. Gregorich, O. Grinstead, C. Gomez, and W. Hussey, *The efficacy of female condom skills training in HIV risk reduction among women: A randomized controlled trial*, Am. J. Public Health 98 (2008), pp. 1841–1848.

[14] L.M. Collins, J.L. Schafer, and C.-M. Kam, *A comparison of inclusive and restrictive strategies in modern missing data procedures*, Psychol. Methods 6 (2001), pp. 330–351.

[15] C. Cox, *Location-scale cumulative odds models for ordinal data: A generalized non-linear model approach*, Stat. Med. 14 (1995), pp. 1191–1203.

[16] R.M. Daniel and M.G. Kenward, *A method for increasing the robustness of multiple imputation*, Comput. Stat. Data Anal. 56 (2012), pp. 1624–1643.

[17] S. Das and R.M. Rahman, *Application of ordinal logistic regression analysis in determining risk factors of child malnutrition in Bangladesh*, Nutr. J. 10 (2011), p. 124.

[18] H. Demirtas and D. Hedeker, *An imputation strategy for incomplete longitudinal ordinal data*, Stat. Med. 27 (2008), pp. 4086–4093.

[19] H. Demirtas, S.A. Freels, and R.M. Yucel, *Plausibility of multivariate normality assumption when multiply imputing non-Gaussian continuous outcomes: a simulation assessment*, J. Stat. Comput. Simul. 78 (2008), pp. 69–84.

[20] A.P. Dempster, N.M. Laird, and D.B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, J. R. Stat. Soc. Ser. B (1977), pp. 1–38.

[21] P.J. Diggle, P. Heagerty, K.Y. Liang, and S.L. Zeger, *Analysis of Longitudinal Data*, Oxford, Oxford University Press, 2002.

[22] P. Diggle and M.G. Kenward, *Informative drop-out in longitudinal data analysis*, Appl. Stat. (1994), pp. 49–93.

[23] M.J. Gameroff, *Using the proportional odds model for health-related outcomes: Why, when, and how with various SAS procedures*. In *SUGI* 30 (2005, April), pp. 205–230.

[24] J.J. Heckman, *The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models*, in *Annals of Economic and Social Measurement*, Vol. 5, NBER, Cambridge, 1976, pp. 475–492.

[25] D.T. Kadengye, W. Cools, E. Ceulemans, and W. Van den Noortgate, *Simple imputation methods versus direct likelihood analysis for missing item scores in multilevel educational data*, Behav. Res. Methods 44 (2012), pp. 516–531.

[26] M.G. Kenward, *Selection models for repeated measurements with non-random dropout: An illustration of sensitivity*, Stat. Med. 17 (1998), pp. 2723–2732.

[27] M.G. Kenward and J. Carpenter, *Multiple imputation: Current perspectives*, Statist. Methods Med. Res. 16 (2007), pp. 199–218.

[28] K.J. Lee and J.B. Carlin, *Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation*, J. Epidemiol. 171 (2010), pp. 624–632.

[29] K.J. Lee, J.C. Galati, J.A. Simpson, and J.B. Carlin, *Comparison of methods for imputing ordinal data using multivariate normal imputation: A case study of non-linear effects in a large cohort study*, Stat. Med. 31 (2012), pp. 4164–4174.

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

[31] I. Liu and A. Agresti, *The analysis of ordered categorical data: An overview and a survey of recent developments*, Test 14 (2005), pp. 1–73.

[32] C.H. Mallinckrodt, S.W.S. Clark, R.J. Carroll, and G. Molenberghs, *Assessing response profiles from incomplete longitudinal clinical trial data under regulatory considerations*, J. Biopharm. Stat. 13 (2003), pp. 179–190.

[33] P. McCullagh, *Regression models for ordinal data*, J. Royal Stat. Soc. Ser. B (1980), pp. 109–142.

[34] X.L. Meng, *Multiple-imputation inferences with uncongenial sources of input*, Statist. Sci. 9 (1994), pp. 538–558.

[35] G. Molenberghs, C. Beunckens, C. Sotto, and M.G. Kenward, *Every missingness not at random model has a missingness at random counterpart with equal fit*, J. R. Statist. Soc. Ser. B 70 (2008), pp. 371–388.

[36] G. Molenberghs and M. Kenward, *Missing Data in Clinical Studies, Vol. 61*, Chichester, Wiley, 2007.

[37] G. Molenberghs and G. Verbeke, *Models for Discrete Longitudinal Data*, Springer, New York, 2005.

[38] J. Newsom, R.N. Jones, and S.M. Hofer (Eds.), *Longitudinal Data Analysis: A Practical Guide for Researchers in Aging, Health, and Social Sciences*, New York, Routledge, 2012.

[39] Y. Pawitan, *In All Likelihood: Statistical Modelling and Inference Using Likelihood*, Oxford, Oxford University Press, 2001.

[40] B. Ratitch, I. Lipkovich, and M. O'kelly, *Combining analysis results from multiply imputed categorical data*, PharmaSUG 2013-PaperSP03, pp. 1–10.

[41] C.H. Rhoads, *Problems with tests of the missingness mechanism in quantitative policy studies*, Stat. Politics Policy 3 (2012).

[42] J.M. Robins and R.D. Gill, *Non-response models for the analysis of non-monotone ignorable missing data*, Stat. Med. 16 (1997), pp. 39–56.

[43] J.M. Robins, A. Rotnitzky, and L.P. Zhao, *Analysis of semiparametric regression models for repeated outcomes in the presence of missing data*, J. Am. Stat. Assoc. 90 (1995), pp. 106–121.

[44] P. Royston, *Multiple imputation of missing values*, Stata J. 4 (2004), pp. 227–241.

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

[46] D.B. Rubin, *Formalizing subjectivenotions about the effect of nonrespondents in sample surveys*, J. Am. Statist. Assoc. 72 (1977), pp. 538–543.

[47] D.B. Rubin, *Multiple Imputation for Nonresponse in Surveys*, Wiley, New York, 1987.
[48] D.B. Rubin and J.L. Schafer, *Efficiently creating multiple imputations for incomplete multivariate normal data*, in *Proceedings of the Statistical Computing Section of the American Statistical Association* 83, American Statistical Association, Alexandria, VA, 1990, p. 88.
[49] J.L. Schafer, *Analysis of Incomplete Multivariate Data*, Chapman and Hall, New York, 1997.
[50] J.L. Schafer, *Multiple imputation in multivariate problems when the imputation and analysis models differ*, Statist. Neerlandica 57 (2003), pp. 19–35.
[51] J.L. Schafer and R.M Yucel, *Computational strategies for multivariate linear mixed-effects models with missing values*, J. Comput. Graph. Stat. 11 (2002), pp. 437–457.
[52] S.R. Seaman, J.W. Bartlett, and I.R. White, *Multiple imputation of missing covariates with non-linear effects and interactions: An evaluation of statistical methods*, BMC Med. Res. Methodol. 12 (2012), p. 46.
[53] R.L. Seitzman, V.B. Mahajan, C. Mangione, J.A. Cauley, K.E. Ensrud, K.L. Stone, S.R. Cummings, M.C. Hochberg, T.A. Hillier, J.S. Sinsheimer, F. Yu, and A.L. Coleman, *Estrogen receptor alpha and matrix metalloproteinase 2 polymorphisms and age-related maculopathy in older women*, Am. J. Epidemiol. 167 (2008), pp. 1217–1225.
[54] M.A. Tanner and W.H. Wong, *The calculation of posterior distributions by data augmentation*, J. Am. Statist. Assoc. 82 (1987), pp. 528–540.
[55] UCLA: Statistical Consulting Group, *Statistical computing seminars. Multiple imputation in STATA, Part1*, from http://www.ats.ucl.edu/stat/stata/seminars/missing_data/mi_in_stat_pt1.htm (accessed on November 12, 2015).
[56] S. Van Buuren, *Multiple imputation of discrete and continuous data by fully conditional specification*, Statist. Methods Med. Res. 16 (2007), pp. 219–242.
[57] S. Van Buuren, *Flexible Imputation of Missing Data*, Boca Raton, FL, CRC Press, 2012.
[58] S. Van Buuren, H.C. Boshuizen, and D.L. Knook, *Multiple imputation of missing blood pressure covariates in survival analysis*, Stat. Med. 18 (1999), pp. 681–694.
[59] S. Van Buuren, J.P.L. Brand, C.G.M. Groothuis-Oudshoorn, and D.B. Rubin, *Fully conditional specification in multivariate imputation*, J. Statist. Comput. Simul. 76 (2006), pp. 1049–1064.
[60] I.R. White, P. Royston, and A.M. Wood, *Multiple imputation using chained equations: Issues and guidance for practice*, Stat. Med. 30 (2011), pp. 377–399.
[61] A.M. Wood, I.R. White, M. Hillsdon, and J. Carpenter, *Comparison of imputation and modelling methods in the analysis of a physical activity trial with missing outcomes*, Int. J. Epidemiol. 34 (2005), pp. 89–99.
[62] L.-M. Yu, A. Burton, and O. Rivero-Arias, *Evaluation of software for multiple imputation of semi-continuous data*, Statist. Methods Med. Res. 16 (2007), pp. 243–258.