Review and Evaluation of Imputation Methods for Multivariate Longitudinal data with Mixed-type Incomplete Variables

Yi Cao[1] │ Heather Allore[2,3] │ Brent Vander Wyk[2] │ Roee Gutman[1]

[1]Department of Biostatistics, Brown University, Providence, RI, USA

[2]Section of Geriatrics, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA

[3]Department of Biostatistics, Yale School of Medicine, New Haven, CT, USA

Correspondence
Yi Cao, Department of Biostatistics, Brown University, 121 South Main Street, Providence, RI 02903l, USA.
Email: yi_cao2@brown.edu

Summary

Estimating relationships between multiple incomplete patient measurements requires methods to cope with missing values. Multiple imputation is one approach to address missing data by filling in plausible values for those that are missing. Multiple imputation procedures can be classified into two broad types: joint modeling (JM) and fully conditional specification (FCS). JM fits a multivariate distribution for the entire set of variables, but it may be complex to define and implement. FCS imputes missing data variable-by-variable from a set of conditional distributions. In many studies, FCS is easier to define and implement than JM, but it may be based on incompatible conditional models. Imputation methods based on multilevel modeling show improved operating characteristics when imputing longitudinal data, but they can be computationally intensive, especially when imputing multiple variables simultaneously. We review current MI methods for incomplete longitudinal data and their implementation on widely accessible software. Using simulated data from the National Health and Aging Trends Study, we compare their performance for monotone and intermittent missing data patterns. Our simulations demonstrate that in a longitudinal study with a limited number of repeated observations and time-varying variables, FCS-Standard is a computationally efficient imputation method that is accurate and precise for univariate single-level and multilevel regression models. When the analyses comprise multivariate multilevel models, FCS-LMM-latent is a statistically valid procedure with overall more accurate estimates, but it requires more intensive computations. Imputation methods based on generalized linear multilevel models can lead to biased subject-level variance estimates when the statistical analyses involve hierarchical models.

KEYWORDS:
longitudinal analysis, multiple imputation, chained equations, joint modeling

1 | INTRODUCTION

Missing data are often inevitable in longitudinal studies. A primary reason is that non-response can occur at any time in the study. Individuals' responses may be missing because they have moved out of the area, missed an appointment, were too ill to attend, or died. In studies involving annual surveys, missing data also occur when participants refuse to answer or do not know the answer.

Commonly used statistically valid methods can be classified into three broad types: (1) likelihood and Bayesian methods; (2) weighting methods; and (3) imputation methods.[1] The likelihood and Bayesian methods define a model for the observed and unobserved variables. Using computational techniques such as the EM algorithm[2] or Data Augmentation,[3] it provides estimates for the estimands of interest. These methods may result in biased estimates when the model is mis-specified or when the missing data mechanism is non-ignorable. Weighting is an alternative approach to handling missing data. This approach weights the observed data to account for missing observations using the estimated probabilities of non-response.[4–6] Weighting methods are best suited for monotone missing data patterns and are commonly used when the missing variable is scalar. In contrast to the former two methods, imputation methods explicitly "fill in" the missing values with plausible values. Single imputations have been shown to result in sampling variance estimates that are too small.[7] Multiple imputation procedures circumvent this issue by replacing each missing value with a set of B plausible values drawn from the predictive distribution of the missing data. Based on these values, B sets of complete datasets are generated. Each dataset is analyzed separately, and final estimates are obtained using common combination rules.[8]

Although the idea behind multiple imputation seems simple, developing procedures to produce plausible values is more complex. Multiple papers have summarized and compared possible procedures to impute scalar variables,[9] multiple non-clustered variables,[10–13] and multiple continuous longitudinal variables.[14–18] However, methods to impute mixed-type variables in longitudinal studies are more limited and dispersed. Kalaycioglu et al[19] compared three chained equations imputation approach,[20] the multivariate Normal imputation,[11] and a Bayesian imputation approach to impute time-varying binary, categorical, skewed, and normally distributed variables. Huque et al[21] presented a comparison study with twelve currently available imputation methods for longitudinal data with incomplete continuous and binary variables. Other studies compared imputation methods for continuous and binary multilevel data,[16,22–25] and multilevel categorical data[26] in non-longitudinal settings. This manuscript reviews available imputation methods for multiple variables of various types in longitudinal studies. Using simulations based on the National Health and Aging Trend Study (NHATS),[27] we compare the operating characteristics of different methods for handling missing values in both outcomes and explanatory variables. Code in R[28] for implementing all of the methods for the simulations and the real-data example is provided.[29]

The paper proceeds as follows. Section 2 introduces the multiple imputation approach. Sections 3 and 4 review the fully conditional specification (FCS) imputation procedures and the joint modeling (JM) imputation procedures, respectively. Section 5 describes the simulation analyses and presents the results of the simulations. An application for estimating the associations between hospital admissions, physical, cognitive abilities, and skilled nursing facilities admissions by different imputation methods is demonstrated in Section 6. Section 7 provides a discussion and conclusions.

2 | MULTIPLE IMPUTATION FOR LONGITUDINAL STUDIES

2.1 | Notations and assumptions

Let $\mathbf{Y} = \{y_{ijk}\}$ represent a multivariate longitudinal dataset such that y_{ijk} is the value of variable $\mathbf{Y}_j$, $j \in \{1, \dots, J\}$, for subject $i \in \{1, \dots, n\}$ at time t_k, $k \in \{1, \dots, K\}$. The data $\mathbf{Y}$ can be saved in either a wide or long matrix format. In wide format, $\mathbf{Y}$ is stored as a $n \times JK$ matrix with each row representing an individual and each column corresponding to variable $\mathbf{Y}_j$ measured at time t_k, denoted by $\mathbf{Y}_{jk}$. In long format, $\mathbf{Y}$ is a $nK \times (J + 1)$ matrix, with the last column describing the time point that the observation is recorded and the other columns are the J variables. For simplicity, we define $\mathbf{Y}_l$, $l \in \{1, \dots, L\}$, to be a column in either the wide or long format. In wide format, $L = JK$ and $\mathbf{Y}_l$ is a vector of size n. In long format, $L = J$ and $\mathbf{Y}_l$ is a vector of size nK. Tables 1 and 2 present an example of data arranged in wide format and long format, respectively, with $K = 4$ time points and $J = 3$ variables. Question marks represent missing values.

Let $\mathbf{M} = \{m_{ijk}\}$ be a matrix of indicators, such that $m_{ijk} = 0$ when y_{ijk} is observed and $m_{ijk} = 1$, otherwise. In addition, let $\mathbf{Y}_l^{obs}$ and $\mathbf{Y}_l^{mis}$ be the observed ($m_{ijk} = 0$) and the missing ($m_{ijk} = 1$) parts of $\mathbf{Y}_l$, respectively. Monotone and intermittent missing data are two missing data patterns that are commonly observed in longitudinal data.[7] Monotone missing data occurs when subjects drop out from the study and do not return for follow-up appointments. If subject i drops out from the study at time point t_k, then $m_{ijk^*} = 1$, $\forall j \in \{1, \dots, J\}$ and $k^* \geq k$. Intermittent missing data usually occur when subjects skip an interview or refuse to answer certain questions. In such case, $m_{ijk} = 1$ for any $j \in \{1, \dots, J\}$ and $k \in \{1, \dots, K\}$. Imputation of $\mathbf{Y}_l^{mis}$ involves modeling the relationships between $\mathbf{Y}_l$, complete covariates $\mathbf{X} = \{x_{ip}\}$, $p \in \{1, \dots, P\}$, and all the other $L - 1$ variables, $\mathbf{Y}_{-l}$.

To obtain valid inferences when data comprise missing values, one should consider the missing data mechanism,[7] $P(\mathbf{M}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\phi})$, where $\boldsymbol{\phi}$ are the parameters governing this distribution. In many software packages, the default method for missing values is the list-wise deletion procedure, which assumes that the data are missing completely at random (MCAR). MCAR implies that the missing data mechanism is unrelated to missing and observed values, $P(\mathbf{M}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\phi}) = P(\mathbf{M}|\boldsymbol{\phi})$. In aging research, older adults may be too ill to complete the study, and assuming MCAR can lead to biased estimates. One way to relax the MCAR assumption is to assume that missing data depends only on observed values, also know as missing at random (MAR), $P(\mathbf{M}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\phi}) = P(\mathbf{M}|\mathbf{Y}^{obs}, \mathbf{X}, \boldsymbol{\phi})$, where $\mathbf{Y}^{obs} = \{\mathbf{Y}_l^{obs}\}$. Under MAR, a variety of analytic strategies to address

missing data can be considered. The third type of missing data mechanism is not missing at random (NMAR). Under NMAR, missing data depends on both the observed and unobserved values. Because missing data depends on unobserved values, methods to handle missing data under NMAR rely on assumptions that are not verifiable from the observed data. Thus, many authors emphasized the need for sensitivity analysis to assess inferences under different plausible assumptions.[1,30] This paper describes multiple imputation methods that assume that the missing data mechanism is MAR, and it examines their performance when these methods are applied to multiple incomplete longitudinal variables of different types.

2.2 | Multiple imputation for multivariate data

Assuming MAR, a multiple imputation (MI) procedure generates B plausible values for each missing value resulting in B complete datasets. Each complete dataset is analyzed separately, and point and interval estimates are obtained using common combination rules. The procedure for creating imputations for all partially observed variables in $\mathbf{Y}$ consists of three steps:

1. Specify the model $P(\mathbf{Y}_1, \dots, \mathbf{Y}_L | \theta, \mathbf{X})$ and prior distributions of parameters $P(\theta)$ and calculate the posterior distribution of θ based on $\mathbf{Y}^{obs}$;

2. Draw a value θ^* from its posterior distribution $P(\theta | \mathbf{Y}^{obs}, \mathbf{X})$;

3. Draw imputations $\mathbf{Y}_1^*, \dots, \mathbf{Y}_L^*$ from the conditional posterior predictive distribution of $\mathbf{Y}^{mis}$ given $\mathbf{Y}^{obs}$, θ^*, and $\mathbf{X}$, $P(\mathbf{Y}^{mis} | \mathbf{Y}^{obs}, \theta^*, \mathbf{X})$.

After the imputation for all incomplete variables is completed, researchers can conduct any statistical analysis that they would have performed on a complete dataset. Using common combination rules, an estimate for a scalar parameter of interest, β, is derived as the average of $\hat{\beta}^{(u)}$, $u \in \{1, \dots, B\}$, which is the estimate of β within complete dataset u. And its sampling variance, $\text{var}(\hat{\beta})$, is estimated by summing the average sampling variances within imputation and the variance between imputations[8]. Formally,

$$\hat{\beta} = \frac{1}{B} \sum_{u=1}^{B} \hat{\beta}^{(u)}, \ \text{var}(\hat{\beta}) = \frac{1}{B} \sum_{u=1}^{B} U_B + (1 + \frac{1}{B}) W_B, \ \text{where } U_B = \sum_{u=1}^{B} \hat{\text{var}}(\beta^{(u)}), \ W_B = \frac{1}{B-1} \sum_{u=1}^{B} (\hat{\beta}^{(u)} - \hat{\beta})^2. \quad (2.1)$$

Two main strategies for specifying a distribution of all incomplete variables have been proposed: joint modeling (JM) and fully conditional specification (FCS).[1] The JM approach defines a multivariate distribution for all incomplete variables, $P(\mathbf{Y}_1, \dots, \mathbf{Y}_L | \theta, \mathbf{X})$. The FCS approach specifies a set of univariate conditional distributions for each incomplete variable given the other variables, $\{P(\mathbf{Y}_l | \theta_l, \mathbf{Y}_{-l}, \mathbf{X})\}$. Compared to the JM approach, the FCS approach provides more flexibility for imputing different types of variables, but it may suffer from theoretical limitations, because the joint distribution based on the different conditional models may not exist.[13]

Under either the FCS or JM strategy, the imputation models used to create $\mathbf{Y}_l^*$ depend on the data format. With long format, multilevel models are commonly used as imputation models, where the variance at the subject level captures the correlation between repeated measurements.[31–33] In wide format data, single-level models are specified to impute an incomplete variable at a specific time point. The correlation between the repeated measurements of an incomplete variable is accounted for by adjusting for its values measured at all the other time points as constant effects.[34] This approach assumes an unstructured correlation structure between the repeated measurements of the incomplete variables. While a wide format can be used when the same observation is recorded for the same unit over time, it may be impractical when the same observation is recorded for different units, but units are grouped within clusters. The imputation methods using wide format data should generally be used for balanced longitudinal data in which information on all individuals is recorded over similar intervals. When individuals' reporting is recorded over different time intervals, special care should be given to the time between reports.

3 | IMPUTATION BY FULLY CONDITIONAL SPECIFICATIONS

The FCS approach is also referred to multivariate imputation by chained equations (MICE). MICE iterates through sampling from the conditional posterior distributions of model parameters, $P(\theta_l | \mathbf{Y}_l^{obs}, \mathbf{Y}_{-l}, \mathbf{X})$, and sampling from the conditional posterior predictive distributions $P(\mathbf{Y}_l^{mis} | \mathbf{Y}_l^{obs}, \mathbf{Y}_{-l}, \mathbf{X}, \theta_l^*)$. Formally, the t-th iteration of MICE involves sampling

$$
\begin{aligned}
\theta_1^{*(t)} &\sim P(\theta_1 | \mathbf{Y}_1^{obs}, \mathbf{Y}_{-1}^{(t-1)}, \mathbf{X}), \\
\mathbf{Y}_1^{*(t)} &\sim P(\mathbf{Y}_1 | \mathbf{Y}_1^{obs}, \mathbf{Y}_{-1}^{(t-1)}, \mathbf{X}, \theta_1^{*(t)}), \\
&\vdots \\
\theta_L^{*(t)} &\sim P(\theta_L | \mathbf{Y}_L^{obs}, \mathbf{Y}_{-L}^{(t)}, \mathbf{X}), \\
\mathbf{Y}_L^{*(t)} &\sim P(\mathbf{Y}_L | \mathbf{Y}_L^{obs}, \mathbf{Y}_{-L}^{(t)}, \mathbf{X}, \theta_L^{*(t)}).
\end{aligned}
\tag{3.1}
$$

Because the conditional models in MICE may not represent a joint distribution, there are no theoretically supported methods to assess convergence. One possible method to assess convergence is by examining whether the imputation of each variable has converged. For example, by examining the convergence of a summary statistic that utilizes the imputed values in each incomplete variable over multiple chains.[35] The MICE algorithm is implemented in the `mice` package[36] in R, where researchers specify a set of univariate imputation models for each incomplete variable. Other available software for implementing the MICE algorithm include IVEware,[37] `PROC MCMC` in `SAS` with FCS statement. In all of these implementations, FCS model specifications commonly depend on the data format.

3.1 | FCS using wide format data

For wide format longitudinal data, imputation models treat repeated measurements of an incomplete variable $\mathbf{Y}_j$ at time t_k as K distinct variables, $\{\mathbf{Y}_{jk}; k \in \{1, ..., K\}\}$. If $\mathbf{Y}_j$ contains missing values at and after time t_k, then variables $\{\mathbf{Y}_{jt}; \forall t \geq k\}$ are incomplete variables. To impute a missing variable $\mathbf{Y}_{jk}$, the imputer specifies a linear regression model or a generalized linear model where $\mathbf{Y}_{jk}$ is the dependent variable, and $\mathbf{X}$, the variables $\{\mathbf{Y}_{jt}; \forall t \neq k\}$ and all the other variables at all time points $\{\mathbf{Y}_{\tilde{j}t}; \forall \tilde{j} \neq j, t \in \{1, ..., K\}\}$ are the independent variables. Another method for imputing incomplete continuous and count variables is the predictive mean matching (PMM) procedure.[4] PMM is a semi-parametric method that imputes data using observed values, making it less sensitive to model mis-specification than purely parametric methods.[38] For count variables with large number of zeros, the zero-inflated Poisson and the zero-inflated Negative-Binomial models can be used.[39] These models are implemented in the `countimp` package.[40] The application of MICE to wide data format has been referred to as FCS-Standard[21] or as imputation by chained equations with fixed effects regression models (ICE-FS).[19]

As the number of waves in longitudinal studies increases, FCS-Standard can result in numerical instabilities because of the lack of identification that arises from specifying many explanatory variables in the conditional models. Nevalainen et al[41] proposed to impute variables recorded at time t_k only with variables that are recorded within $(t_k - \delta, t_k + \delta)$ time window. This procedure assumes that a partially recorded variable observed at time t_k is independent from variables recorded at time $t_k \pm (\delta + \epsilon)$ $(\epsilon > 0)$ conditional on variables recorded at time $t_k \pm \delta$. This reduces the number of covariates used within each conditional model. Additional details of this FCS method are provided in Welch et al.[42]

3.2 | FCS– Multilevel linear model

When data are saved in long format, multilevel linear models have been proposed as conditional imputation models for continuous variables. The first level of the model describes the repeated observations of subjects across time, and it is nested within a second-level, which is subject-level information. Formally, to impute the missing observation of a continuous variable $\mathbf{Y}_l$ at time point k for the i-th subject, the following multilevel linear model is used,

$$y_{ilk} = \mathbf{X}_{ik}^T \boldsymbol{\beta}_l + \mathbf{Z}_{ik}^T \mathbf{b}_i + e_i, \quad \mathbf{b}_i \sim N_q(0, V_b), \quad e_i \sim N(0, \sigma_e^2), \tag{3.2}$$

where $\mathbf{X}_{ik}$ and $\mathbf{Z}_{ik}$ are $p \times 1$ vector of covariates and $q \times 1$ vector of subject-level covariates, respectively, and V_b is an unstructured covariance matrix of the subject-level effects. Both $\mathbf{X}$ and $\mathbf{Z}$ may comprise complete and other incomplete variables. Parameters $\boldsymbol{\beta}_l$ are regression coefficients corresponding to covariates $\mathbf{X}$, and $\mathbf{b}_i$ correspond to subject-level variations in covariates $\mathbf{Z}$. Model (3.2) assumes a common conditional variance σ_e^2 across all subjects. Applying MICE with Model (3.2) involves sampling $\theta_l^* = (\boldsymbol{\beta}_l^*, V_b^*, \sigma_e^{2*})$, subject-level effects $\mathbf{b}_i^*$, and imputation values y_{ilk}^*. Commonly, conjugate prior distributions are assumed for θ_l^*. With continuous incomplete variables, a multivariate Gaussian prior distribution for $\boldsymbol{\beta}_l$, an Inverse-Wishart distribution

for V_b, and an Inverse-Gamma distribution for σ_e^2 are assumed. Samples of $\mathbf{b}_i$ can be drawn from its conditional posterior distribution $N(\mu_{bi}, \Psi_{bi})$, with $\mu_{bi} = V_b \mathbf{Z}_i^T (\mathbf{Z}_i V_b \mathbf{Z}_i^T + \sigma_e^2 \mathbf{I}_K)^{-1} (\mathbf{y}_{il} - \mathbf{X}_i^T \beta_l)$ and $\Psi_{bi} = V_b - V_b \mathbf{Z}_i^T (\mathbf{Z}_i V_b \mathbf{Z}_i^T + \sigma_e^2 \mathbf{I}_K)^{-1} \mathbf{Z}_i V_b$, where $\mathbf{y}_{il} = (y_{il1}, ..., y_{ilK})^T$ is the responses of subject i measured at all K time points and $\mathbf{I}_K$ denotes a $K \times K$ identity matrix. Imputed values of y_{ilk}^* are sampled from Model (3.2) given θ_l^* and $\mathbf{b}_i^*$. The sampling procedure relies on a MCMC algorithm, which can be computationally burdensome because many samples are required for the chain to converge to its equilibrium distribution.[43] A possible approximation procedure samples θ_l^* from their large sample Normal approximation and $\mathbf{b}_i^*$ from $N(\mu_{b_i}, \Psi_{b_i})$.[16,25]

To impute discrete variables, one approach is to sample from multilevel linear models and round the imputed continuous values to the nearest valid discrete values. However, this rounding step can result in biased estimates.[44] To address this, Yucel et al[45] have proposed a calibration method that is similar to posterior predictive checks in Bayesian analysis[46] to improve imputed rounded values.

3.3 | FCS– Multilevel linear model with latent variables

Another approach for imputing binary and categorical variables is to sample from a multilevel linear model with latent variables.[47,48] For a binary variable $\mathbf{Y}_l$, this model assumes that there is a latent Normal variable $\tilde{\mathbf{Y}}_l$, such that $y_{ilk} = 1$ if $\tilde{y}_{ilk} > 0$, and $y_{ilk} = 0$ otherwise. The latent variable is assumed to follow a multilevel linear model. This representation is equivalent to assuming that y_{ilk} follows a multilevel probit model. Formally,

$$\Phi^{-1}(P(y_{ilk} = 1 | \mathbf{X}_{ik}, \mathbf{b}_i, \theta_l, \tau)) = \tilde{y}_{ilk} = \mathbf{X}_{ik}^T \beta_l + \mathbf{Z}_{ik}^T \mathbf{b}_i + e_i, \quad \mathbf{b}_i \sim N_q(0, V_b), \quad e_i \sim N(0, 1), \tag{3.3}$$

where Φ^{-1} is the inverse cumulative distribution function (CDF) of the standard Normal distribution. The sampling procedure for drawing imputations from Model (3.3) is the same as the procedure for Model (3.2), which can be implemented by the function `mice.impute.2l.jomo` in the `micemd` R package.

For ordinal categorical variables $\mathbf{Y}_l$ with $H > 2$ levels, the latent variable imputation model is based on a cumulative probit model. The model assumes that $\mathbf{Y}_l$ is determined by a latent Normal variable $\tilde{\mathbf{Y}}_l$ partitioned by $H - 1$ threshold parameters $\tau = \{\tau_h\}$, $h \in \{1, ..., H\}$, such that $y_{il} = h$, if $\tau_{h-1} < \tilde{y}_{il} < \tau_h$ ($\tau_0 = -\infty$, $\tau_H = \infty$). In addition, the model assumes that the latent variable $\tilde{Y}_l$ follows a multilevel linear model as Model (3.3). Formally, a cumulative probit model of $\mathbf{Y}_l$ is defined as

$$\Phi^{-1}(P(y_{ilk} \leq h | \mathbf{X}_{ik}, \mathbf{b}_i, \theta_l)) = \Phi^{-1}(P(\tilde{y}_{ilk} < \tau_h | \mathbf{X}_{ik}, \mathbf{b}_i, \theta_l)) = \tau_h - (\mathbf{X}_{ik}^T \beta_l + \mathbf{Z}_{ik}^T \mathbf{b}_i). \tag{3.4}$$

The sampling procedure at iteration t starts with updating $\tau^{(t)} \sim N(\tau | \mathbf{Y}_l, \tilde{\mathbf{Y}}_l^{(t-1)}, \theta_l^{(t-1)}, \mathbf{b}_i^{(t-1)})$ followed by sampling $\tilde{\mathbf{Y}}_l \sim TN(\tilde{\mathbf{Y}}_l | \mathbf{Y}, \tau^{(t)}, \theta_l^{(t-1)}, \mathbf{b}_i^{(t-1)})$ (a truncated Normal distribution) for all subjects, where $\tau^{(t)}$ are the truncation parameters at iteration t, and $\theta_l^{(t-1)} = (\beta_l^{(t-1)}, V_b^{(t-1)})^T$ and $\mathbf{b}_i^{(t-1)}$ are the sample of $\theta_l = (\beta_l, V_b)^T$ and $\mathbf{b}_i$ at iteration $t - 1$, respectively. Samples of the parameters $\theta_l^{(t)}$ are drawn from its conditional posterior distribution $P(\theta_l | \mathbf{Y}_l, \tilde{\mathbf{Y}}_l^{(t)}, \tau^{(t)}, \mathbf{b}_i^{(t-1)})$ and $\mathbf{b}_i^{(t)}$ are drawn from

$P(\mathbf{b}_i|\mathbf{Y}_l, \tilde{\mathbf{Y}}_l^{(t)}, \boldsymbol{\tau}^{(t)}, \boldsymbol{\theta}_l^{(t)})$. The conjugate prior distributions that are used for Model (3.3) are commonly specified for $\boldsymbol{\theta}_l$ and $\mathbf{b}_i$. Additional technical details are provided in Enders et al.[48]

The multilevel multinomial probit model[49] can be used to impute a nominal categorical variable $\mathbf{Y}_l$ with H_l categories. This model expands $\mathbf{Y}_l$ into H_l binary variables $\mathbf{Y}_l^h$, $h \in \{1, ..., H_l\}$, that indicate whether $y_{il} = h$ for subject i. An underlying latent Normal variable $\tilde{\mathbf{Y}}_l^h$ corresponding to $\mathbf{Y}_l^h$ is defined by the probability of $y_{il}^h = 1$. If $\tilde{y}_{il}^h$ is greater than $\tilde{y}_{il}^{h*}$ for all $h^* \neq h$, then $y_{il}^h = 1$ and $y_{il} = h$. For identifiability purposes, a multivariate linear multilevel model is assumed for the first $H_l - 1$ latent variables, such that $\mathbf{b}_i \sim MVN(0, V_b)$ and the within-subject variance Σ_e is the identity matrix. The sampling procedure is similar to Model (3.3), except that $\tilde{\mathbf{Y}}_l^h$ are generated by an accept-reject algorithm.[47] These models for categorical variables using latent Normal variables are implemented in the software Blimp.[50] We refer to the multilevel linear models with latent variables as FCS-LMM-latent.

3.4 | FCS– Multilevel generalized linear model

The multilevel generalized linear models are a flexible approach to model skewed or non-normally distributed variables. These models are also commonly referred to as generalized linear mixed models (GLMM). Assuming that an incomplete variable $\mathbf{Y}_l$ conditional on item-level covariates $\mathbf{X}$ and $\mathbf{Z}$ has an exponential family probability density function or probability mass function. A GLMM is defined as

$$g(E(y_{ilk}|\mathbf{X}_{ik}, \mathbf{Z}_{ik}, \mathbf{b}_i, \boldsymbol{\theta}_l)) = \mathbf{X}_{ik}^T\boldsymbol{\beta} + \mathbf{Z}_{ik}^T\mathbf{b}_i, \quad \mathbf{b}_i \sim N_q(0, V_b), \tag{3.5}$$

where $g(\cdot)$ is a function linking the expected value of response y_{ilk} to linear predictors. Sampling from the posterior distribution of $P(\boldsymbol{\theta}_l|\mathbf{X}, \mathbf{Z}, \mathbf{Y}_l)$ can be implemented using MCMC. The latent multilevel variable model in Section 3.3 can be viewed as Model (3.5) with a probit link function. The probit model can perform well for binary and categorical variable, but may not be suitable for skewed or count variables. A different commonly used link function is the logit link for binary or the log-link for count variables. Compared to the probit link, sampling from $P(\mathbf{b}_i|\mathbf{y}_i, \mathbf{X}_i, \boldsymbol{\theta}_l)$ with the logit or log link functions can be more complex. A possible approximation can be obtained by sampling $\mathbf{b}_i$ from its marginal posterior distribution $N(0, \Psi_b)$, where Ψ_b is estimated by the REML[23] or the Fisher scoring method.[16] Using GLMM models for imputation of binary and count missing variables are implemented in the `micemd` package in R. Availability of software that implements multilevel generalized linear models using the log-Normal or Gamma likelihoods for incomplete skewed continuous data are limited. Throughout, we refer to this method as FCS-GLMM.

4 | IMPUTATION BY JOINT MODELING

4.1 | JM– General location model

The JM approach specifies a multivariate distribution for all incomplete variables. A multivariate Normal distribution is often used when the data are arranged in wide format and consist of only continuous variables.[14,17] For a mixture of continuous and discrete variables, the general location model is proposed as a possible imputation model.[11] This model describes the joint distribution of $\mathbf{Y} = (\mathbf{W}, \mathbf{C})$ in terms of a marginal distribution for all discrete variables, $\mathbf{W} = (\mathbf{W}_1, ..., \mathbf{W}_{S_1})$, and a conditional distribution of all continuous variables $\mathbf{C} = (\mathbf{C}_1, ..., \mathbf{C}_{S_2})$ given the discrete variables, where $S_1 + S_2 = L(= JK)$. The general location model is defined as

$$P(\mathbf{W}, \mathbf{C}) = P(\mathbf{W})P(\mathbf{C}|\mathbf{W})$$
$$= Multinomial(N, \pi_d) \cdot Normal(\mu_d, \Sigma).$$

(4.1)

The marginal distribution $P(\mathbf{W})$ of S_1 discrete variables is modeled by a multinomial distribution on the cell counts of a S_1-dimensional contingency table with $D = \prod_{l=1}^{S_1} H_l$ cells and cell probabilities π_d, $d \in \{1, .., D\}$, where H_l is the number of distinct levels of variable $\mathbf{W}_l$. Within each cell of the contingency table, continuous variables $\mathbf{C}$ follow a multivariate Normal distribution with mean μ_d and covariance matrix Σ_d. In finite samples, as the number of categorical variables increases, some cells may be empty. This may lead to unstable estimation.[51] In these situations, the restricted general location model can be applied. The restricted model assumes that a contingency table cell counts follow a log-linear model, which is fitted by a subset of $\mathbf{W}_l$, $l \in \{1, ..., S_1\}$, and possibly their interactions. Continuous variables are modeled by a multivariate linear regression model with the categorical variables as the independent variables. Another possible limitation of both the general location model and the restricted general location model is their reliance on the multivariate Normal distributions for continuous $\mathbf{Y}_l$. This may result in inaccurate imputation when $\mathbf{Y}_l$ is skewed or multi-modal. The general location model and the restricted location model are implemented in the `mix` package[52] in R. Throughout, we refer to the general location model as JM-GL.

4.2 | JM– Multivariate multilevel linear model

When $\mathbf{Y}$ comprise only continuous variables and is in a long format, a possible joint model for imputing $\mathbf{Y}$ is the multivariate linear multilevel model. Let $\mathbf{y}_{ik} = (y_{i1k}, ..., y_{iLk})^T$ be a column vector of L continuous responses of the i-th subject measured at time point k. The multivariate multilevel linear model (MLMM) is

$$\mathbf{y}_{ik} = (\mathbf{I}_L \otimes \mathbf{X}_{ik}^T)\boldsymbol{\beta} + (\mathbf{I}_L \otimes \mathbf{Z}_{ik}^T)\mathbf{b}_i + \mathbf{e}_i,$$
$$\mathbf{b}_i \sim N(0, V_b), \quad \mathbf{e}_i \sim N(0, \Sigma_e),$$

(4.2)

where the column vector $\boldsymbol{\beta}$ has pL elements and column vector $\mathbf{b}_i$ has qL elements. The symbol $\otimes$ is the Kronecker product. Schafer and Yucel[17] proposed a MCMC procedure to sample $\mathbf{b}_i$, $\boldsymbol{\theta}$, and $\mathbf{Y}^{mis}$ jointly. It assumed that $\Sigma \sim \text{Inv-Wishart}(\nu_1, \Lambda_1)$, $V_b \sim \text{Inv-Wishart}(\nu_2, \Lambda_2)$, and an improper uniform density over $\mathcal{R}^{PL}$ for $\boldsymbol{\beta}$. This model is implemented in R package pan.[53]

4.3 | JM-Multivariate multilevel linear model with latent variables

For imputations of both continuous and categorical variables, a multivariate multilevel linear model (MLMM) can be used to model latent Normal variables that correspond to each of the categorical variables together with other continuous variables. Formally, a MLMM assumes that a set of incomplete continuous variables, $\mathbf{y}_{ik}^{(c)} = (y_{i1k}^{(c)}, ..., y_{iCk}^{(c)})^T$, and a set of latent variables $\tilde{\mathbf{y}}_{ik}^{(w)} = (\tilde{y}_{i1k}^{(w)}, ..., \tilde{y}_{iWk}^{(w)})^T$ of categorical variables $\mathbf{y}_{ik}^{(w)} = (y_{i1k}^{(w)}, ..., y_{iWk}^{(w)})^T$ are distributed as

$$\mathbf{y}_{ik}^{(c)} = (\mathbf{I}_C \otimes \mathbf{X}_{ik}^T)\boldsymbol{\beta}_c + (\mathbf{I}_C \otimes \mathbf{Z}_{ik}^T)\mathbf{b}_{ci} + \mathbf{e}_{ci},$$

$$\tilde{\mathbf{y}}_{ik}^{(w)} = \Phi^{-1}(P(\mathbf{y}_{ik}^{(w)} = 1)) = (\mathbf{I} \otimes \mathbf{X}_{ik}^T)\boldsymbol{\beta}_w + (\mathbf{I}_W \otimes \mathbf{Z}_{ik}^T)\mathbf{b}_{wi} + \mathbf{e}_{wi},$$

$$\mathbf{b}_i = (\mathbf{b}_{ci}, \mathbf{b}_{wi})^T \sim N(0, V_b), \mathbf{e}_i = (\mathbf{e}_{ci}, \mathbf{e}_{wi})^T \sim N(0, \Sigma_e), \text{ where } \Sigma_e = \begin{pmatrix} \sigma_e^2 \mathbf{I}_C & cov(\mathbf{e}_c, \mathbf{e}_w) \\ cov(\mathbf{e}_c, \mathbf{e}_w) & \mathbf{I}_W \end{pmatrix}.$$

$$(4.3)$$

Under this model, a vector stacking the latent variables $\tilde{\mathbf{y}}^{(w)}$ and the continuous variables $\mathbf{y}^{(c)}$ follows a multivariate Normal distribution, where V_b is an unstructured covariance matrix of the subject-level effects, and the covariance matrix Σ_e captures the associations between the two sets of variables. The imputation algorithm is similar to sampling from the Model (4.2). However, Σ_e can not be sampled from the Inverse-Wishart distribution. Instead, the elements of Σ_e should be updated individually using a Metropolis-Hastings procedure. Detailed descriptions of the imputation algorithm are provided in Carpenter and Kenward[54] chapters 4-5. We refer to this model as JM-MLMM-latent and it is implemented in the jomo package in R[55] and the REALCOM program in MATLAB.[56]

4.4 | JM– Multivariate generalized multilevel linear model

Extending the FCS-GLMM to its JM version is an another approach for handling mixed-type incomplete variables. Let $\mathbf{y}_i = \{\mathbf{y}_{i1}, ..., \mathbf{y}_{iL}\}$ be a $K \times L$ response matrix of subject i, which consists of different types of variables. We assume a multivariate generalized linear mixed model (MGLMM) for $P(\mathbf{Y}_1, ..., \mathbf{Y}_l|\boldsymbol{\theta})$ defined as

$$p(\mathbf{y}_{i1}, ..., \mathbf{y}_{iL}|\boldsymbol{\theta}) = \int \prod_{l=1}^{L} p_l(\mathbf{y}_{il}|\mathbf{b}_{il}, \boldsymbol{\theta}_l)p_b(\mathbf{b}_i|V_b)d\mathbf{b}_i, \tag{4.4}$$

where $p_l(\cdot)$ are density functions, $\mathbf{b}_i = (\mathbf{b}_{i1}, ..., \mathbf{b}_{iL})^T$ is a vector of subject-level effects which follows a multivariate Normal distribution with mean zero and covariance matrix V_b. The Model (4.4) links a set of univariate generalized linear multilevel models by introducing correlations between the variance components of the subject-level effects. The model assumes that the

Y_ls are independent conditional on $\mathbf{b}_i$ and $\mathbf{X}$. We refer to this model as JM-MGLMM. For example, a shared-random intercepts model with two outcomes ($L = 2$) in Equation (4.4) is

$$E(y_{i1k}|\mathbf{b}_{i0}, \mathbf{X}_{ik}, \beta_{01}) = g_1^{-1}(\beta_{01} + b_{i10} + \beta_1\mathbf{X}_{ik}),$$

$$E(y_{i2k}|\mathbf{b}_{i0}, \mathbf{X}_{ik}, \beta_{02}) = g_2^{-1}(\beta_{02} + b_{i20} + \beta_2\mathbf{X}_{ik}),$$

$$\mathbf{b}_{i0} = (b_{i10}, b_{i20}) \sim N_2(\mathbf{0}, V_b), V_b = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \tag{4.5}$$

where g_1 and g_2 are link functions for outcomes $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively, and the latent correlation of outcomes at the subject-level is identified by ρ. A possible extension of Model (4.5) involves adding random slopes in $E(y_{i1k}|\mathbf{b}_{i0})$ and $E(y_{i2k}|\mathbf{b}_{i0})$. Including all of the covariates as subject-level random slopes in $\mathbf{b}_i = (b_{i10}, ..., b_{i1p}, ..., b_{iL0}, ..., b_{iLp})$ would increase the dimension of V_b to $(L \times (1 + p))^2$. To complete the Bayesian model, diffused prior distributions can be used. Specifically, $\beta_p \sim N(0, 100)$, $V_b \sim$ Inverse-Wishart($\mathbf{I}_q, q$) where q is the cardinality of $\mathbf{b}_i$. Sampling from posterior distributions of the parameters can be implemented using the JAGS software[57], which requires the users to specify both the prior distributions and the likelihood functions. We have provided a code example on the GitHub website.

5 | SIMULATIONS

The National Health and Aging Trend Study (NHATS) collects information on a nationally representative sample of Medicare beneficiaries ages 65 and older. Beginning in 2011, annual interviews are conducted and detailed information on a broad range of variables related to sociodemographic factors, physical, cognitive capacity, and health outcomes are collected. We use the NHATS data collected from 2011 to 2014 (rounds 1-4) and select a set of variables including four incomplete longitudinal variables of varying types: an indicator of whether a person had an overnight hospital stay, a person's body mass index (BMI), comorbidity index, and the count of devices paid to assist with daily activities during the past year (paid assistive devices). The comorbidity index is defined as a count of 10 chronic conditions, including heart disease (e.g., angina or congestive heart failure), hypertension, arthritis, osteoporosis, diabetes, lung disease, Alzheimer's disease, and related dementia, cancer, and whether they experienced a heart attack or a stroke in the past. The number of paid assistive devices (ranges from 0-9) includes the following aids: vision aids, hearing aid, cane, walker, wheelchair, scooters, grabbers, special dress items, and adapted utensils. Almost 98% of the missing values resulted from loss to follow-up. Patients' information collected at the first interview is used as covariates for imputation and analysis, including age, gender, self-rated health, an indicator for whether participants take prescribed medicines, and an indicator for whether participants or their spouse/partner have any medical bills that are being paid off over time. Our goal is to estimate the associations between the comorbidity index, BMI, the paid assistive devices, and the hospital stay status after adjustment for baseline variables.

We design a simulation study to evaluate the operating characteristics of different imputation methods described in Sections 3 and 4. Our simulations are based on 5309 participants that were observed for 4 all rounds, and we simulate different missing data patterns.

5.1 | Missing data mechanism and missing data patterns

In the simulation, we assume that the missing data mechanism is MAR for the monotone missing data pattern. In longitudinal studies with many covariates, it is reasonable to believe that $P(m_{ijk}|\mathbf{X}_i, Y_{i1}, \dots, Y_{iL})$ would depend on at least one missing Y_{iL} value for some i, where $\mathbf{Y}$ is recorded in a wide format. Thus, for intermittent missing data, we assume an NMAR missing data mechanism. We simulate both missing data patterns on the wide-formatted NHATS data composed of only the completed cases. Missing data indicators are sampled from the Bernoulli distribution with the event probabilities predicted by models estimated from the original NHATS data.

To simulate an intermittent missing data pattern, we generate the missing indicators $m_{ijk} = 1$ for each of the incomplete variables independently. For subject i at round $k \in \{2, 3, 4\}$, the probability of Y_{ijk} being missing is

$$P(m_{ijk} = 1|\mathbf{Y}_i, \mathbf{X}_i) = \text{logit}^{-1}(\alpha_{0jk}^* + \sum_{j=1}^{J} \sum_{k=1}^{k-1} \hat{\alpha}_{jk}^* y_{ijk} + \sum_{p=1}^{P} \hat{\psi}_{jp} x_{ip}), \tag{5.1}$$

where the covariates include all of the time-varying variables prior to round k and time-invariant variables. We set α_{0jk}^* to ensure that the average missing proportions of $\mathbf{Y}_{jk}$ at round 2, 3, and 4 are 20%, 35%, and 40%, respectively. The coefficients $\hat{\alpha}_{jk}^*$ and $\hat{\psi}_{jp}$ are the maximum likelihood estimates of Model (5.1) using the original NHATS data.

To simulate a monotone missing data pattern, we generate the drop-out indicators for each subject. Let r_{ik} represent a drop-out indicator for subject i at round $k \in \{2, 3, 4\}$. If subject i drops out at round k^*, then $r_{ik^*} = 1$ and values for subject i are not observed at round k^* and in subsequent rounds for all variables ($m_{ijk} = 1, \forall k \geq k^*$ and $\forall j \in \{1, ..., J\}$). The probability that subject i is lost to follow-up at round k is

$$P(r_{ik} = 1|\mathbf{Y}_i^{obs}, \mathbf{X}_i) = \text{logit}^{-1}(\tilde{\alpha}_{0k} + \sum_{j=1}^{J} \sum_{k=1}^{k-1} \hat{\alpha}_{jk} y_{ijk} + \sum_{p=1}^{P} \hat{\psi}_p x_{ip}), \tag{5.2}$$

where the covariates include all of the time-varying variables that are fully observed prior to round k and time-invariant variables. The intercept $\tilde{\alpha}_{0k}$ is set to ensure a pre-specified proportion of participants who drop out. Based on the original NHATS data, we set the proportion of participants who start to drop out from the study at rounds 2, 3, and 4 to 20%, 15% and 10%, respectively. Cumulatively, the proportion of individuals with missing information at round 4 is approximately 45%. The coefficients $\hat{\alpha}_{jk}$ and $\hat{\psi}_p$ are the maximum likelihood estimates of Model (5.2) using the drop-out indicators observed in NHATS.

5.2 | Study design

We consider three configurations with different numbers and types of incomplete variables. The variables include a binary hospital stay status ($\mathbf{Y}_1$), a discrete bounded comorbidity index ($\mathbf{Y}_2$), a continuous BMI ($\mathbf{Y}_3$), and a discrete count number of paid assistive devices ($\mathbf{Y}_4$). Configuration 1 assumes that $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are incomplete and the other two variables are fully observed. Configuration 2 assumes that in addition to $\mathbf{Y}_1$ and $\mathbf{Y}_2$, $\mathbf{Y}_3$ is incomplete. Lastly, Configuration 3 assumes that all four variables are incomplete. Because monotone missing data pattern is observed for most of the individuals in NHATS, we evaluate the performance of the different imputation methods using 500 simulated incomplete datasets with monotone missing data patterns for each of the three configurations. We also compare the different methods on 300 incomplete simulated datasets with intermittent missing data patterns for Configuration 3.

For every simulated dataset, we conduct a multiple imputation procedure with $B = 5$ imputations using 10 imputation methods with different choices of imputation models described in Section 3 and 4. For the FCS-Standard method, we consider three imputation models to impute the count variables: linear regressions, predictive mean matching, and Poisson regressions. When applying FCS-LMM-latent, we use latent variable models for binary variables and linear models for continuous and count variables. For the FCS-GLMM method, we specify a logit link for binary variables, an identity link for continuous variables, and either an identity or a log link for count variables. When applying JM-MLMM-latent, we consider both homoscedastic and heteroscedastic within-subject variance. For JM-MGLMM, we assume a logit link for binary variables, an identity link for continuous variables, and a log link for count variables. A complete list of the different methods is provided in Table 3.

Imputing the missing data is usually performed as part of the data preparation process, and the ultimate goal is to generate unbiased estimates of the associations and conditional associations between variables. Our simulations mimic situations in which no specific statistical analysis is specified prior to the imputation, and the imputed datasets are used for multiple analyses. We conduct three types of analyses: univariate generalized hierarchical model, latent growth model, and bi-variate generalized hierarchical model. An overview of the simulation is presented in Table 4.

5.2.1 | Univariate generalized linear hierarchical model

Within each imputed dataset, a multilevel logistic regression model is used to model the conditional associations between the comorbidity index and the hospital stay status,

$$\text{logit}(P(y_{i1k} = 1 | \boldsymbol{\gamma}, b_{0i}, \mathbf{X}_i, t_{ik}, y_{i2k}, y_{i3k}, y_{i4k})) = \gamma_0 + b_{0i} + \gamma_1 y_{i2k} + \gamma_2 y_{i3k} + \gamma_3 y_{i4k} + \gamma_4 t_{ik} + \mathbf{X}_i^T \boldsymbol{\gamma}_p, \tag{5.3}$$

where $\mathbf{X}_i$ comprises of five baseline covariates for subject i, t_{ik} is the round of time that individual i is being interviewed, $b_{0i} \sim N(0, \sigma_b^2)$ denotes the subject-level effects, and γ_l, $l \in \{0, ..., 8\}$, denotes a set of unknown coefficients. Let $\hat{\gamma}_1^{(u)}, \hat{\gamma}_2^{(u)}, \hat{\gamma}_3^{(u)}$ be

the estimates of Model (5.3) within imputed data set $u = \{1, \dots, 5\}$. and $\hat{\sigma}_{\gamma_1}^2, \hat{\sigma}_{\gamma_2}^2, \hat{\sigma}_{\gamma_3}^2$ be their corresponding sampling variances. The final estimates are obtained using the common combination rules described in Section 2.

5.2.2 | Latent growth curve model

Researchers may be interested in the trajectory of individuals over time. For the imputed datasets in Configuration 3, we fit a latent linear growth curve model (LGCM) on the trajectory of BMI over time. The model includes all the time-invariant variables and the time-varying comorbidity index $\mathbf{Y}_2$ as the predictors. The LGCM requires the data to be structured in wide format, and it can be broken down into two latent constructs, the intercept factor and the slope factor.[58] Let $t_i \in \{0, 1, 2, 3\}$ denote the round that subject i is observed. The adjusted LGCM is expressed by a multilevel model that consists of an intercept model, π_{0i} and a slope model, π_{1i},

$$
\begin{aligned}
y_{i3t} &= \pi_{0i} + \pi_{1i}t_i + \epsilon_{it}, \\
\pi_{0i} &= \eta_{00} + \eta_{01}y_{i2t} + \mathbf{X}_i^T \boldsymbol{\eta}_0 + e_{0i}, \\
\pi_{1i} &= \eta_{10} + \eta_{11}y_{i2t} + \mathbf{X}_i^T \boldsymbol{\eta}_1 + e_{1i},
\end{aligned}
\tag{5.4}
$$

where the residuals of the intercept and slope models are $(e_{0i}, e_{1i})^T \sim N(\mathbf{0}_2, \Sigma)$, and Σ is an unknown unstructured covariance matrix.

5.2.3 | Bivariate generalized linear hierarchical model

In some studies, researchers are interested in examining the associations between multiple factors and multiple outcomes simultaneously.[59–61] This can be achieved by jointly modeling multiple outcomes. Compared to a univariate model, joint models are computationally more complex. We examine a joint model comprising of hospital stay ($\mathbf{Y}_1$) and paid assistive devices ($\mathbf{Y}_4$) on datasets with imputed comorbidity index ($\mathbf{Y}_2$) and BMI ($\mathbf{Y}_3$) for Configuration 3. The bivariate multilevel generalized linear model used in the analysis is

$$
\begin{aligned}
E(y_{i1k}|\mathbf{b}_{i0}, y_{i2k}, y_{i3k}, \mathbf{X}_i, \lambda_{10}, \lambda_{11}, \lambda_{12}) &= \text{logit}^{-1}(\lambda_{10} + b_{i10} + \lambda_{11}y_{i2k} + \lambda_{12}y_{i3k} + \lambda_1\mathbf{X}_{ik}), \\
E(y_{i4k}|\mathbf{b}_{i0}, y_{i2k}, y_{i3k}, \mathbf{X}_i, \lambda_{20}, \lambda_{21}, \lambda_{22}) &= \exp(\lambda_{20} + b_{i20} + \lambda_{21}y_{i2k} + \lambda_{22}y_{i3k} + \lambda_2\mathbf{X}_{ik}), \\
\text{where } \mathbf{b}_{i0} = (b_{i10}, b_{i20})^T \sim N_2(\mathbf{0}, \Sigma_b), \Sigma_b &= \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},
\end{aligned}
\tag{5.5}
$$

where a logit link function is applied to hospital stay and the log link function to the paid assistive devices. The subject-level effect $b_{i0} = (b_{i10}, b_{i20})^T$ is assumed to follow a bivariate Normal distribution with zero means and an unstructured covariance matrix Σ_b. The correlation between the separate random intercepts ρ represents the interdependence between the two outcomes at the subject-level.

5.2.4 | Congeniality of Analysis Models

Congeniality [62] between the different imputation methods and the three analyses varies. Not all of the imputation procedures had the same or more general model specification compared to the three analysis models. For the univariate GLMM, methods that include multilevel modeling encompass the analysis model, whereas methods based on the wide format data are mis-specified. For the bivariate GLMM analysis, JM-MLMM and JM-MGLMM methods are the only two methods that encompass the analysis model. For the latent growth curve, all imputation methods are mis-specified because time was not adjusted for in any of the imputation models.

5.2.5 | Performance assessment metrics

Estimates of the three analyses are obtained using the common combination rules described in Section 2. For each configuration and each replication we estimate the relative bias $(\hat{\theta}^{meth} - \hat{\theta}^{comp})/\hat{\theta}^{comp}$, where $\hat{\theta}^{meth}$ is the estimate obtained after implementation of multiple imputation procedure *meth* described in Table 3, and $\hat{\theta}^{comp}$ is the estimate from the complete dataset. We record the root of mean squared error (RMSE), the 95% interval estimate width, and whether the interval estimate covers the estimate with complete data. Additionally, we estimate the fraction of missing information (FMI) that measures the uncertainty in the imputed values for missing elements. [8,63] For each parameter estimate, the FMI estimate is

$$\hat{\lambda}_m = \frac{r_m + 2/(v+3)}{r_m + 1}, \tag{5.6}$$

where $r_m = (1 + m^{-1})\frac{W_B}{U_B}$, $v = (m-1)(1 + \frac{1}{r})^2$. W_B and U_B are calculated using Equation (2.1), and represent the variance between the m complete-data estimates and the average of the m complete-data variances, respectively. We summarize these metrics by averaging across all replications.

5.3 | Results of the univariate GLMM across all configurations

5.3.1 | Relative bias of coefficients and the subject-level variance

Figure 1 presents the relative bias of regression coefficients estimates associated with the incomplete variables and the subject-level variance estimate, σ_b^2, for study design Configurations 1-3 with monotone missing data patterns. The first row depicts the change in relative bias of $\hat{\gamma}_1$, the conditional log-odds ratio of having a hospital stay with one unit increase in the comorbidity index, as the number of incomplete variables increases from two to four. The boxplots of relative biases are similar for each method as the number of incomplete predictors increases. Across Configurations 1-3, JM-GL results in the smallest mean relative bias of -0.02. However, it has the largest variability in relative biases of $\hat{\lambda}_1$ compared to other methods for all configurations. FCS-Standard with either linear regressions (LM) or with predictive mean matching (PMM), and FCS-LMM-latent result in comparable mean relative biases to JM-GL. Across Configurations 1-3, the averages relative bias for LM is between -0.04

to -0.03, and -0.03 to -0.02 for PMM and FCS-LMM-latent. Imputation methods that assume Poisson regressions for the count variables, FCS-Standard, FCS-GLMM, and JM-MGLMM have mean relative biases that are greater than -0.29. The JM-MLMM methods generally have the second largest relative bias. Across Configurations 1-3, JM-MGLMM leads to the largest mean relative bias of approximately -0.40.

Similar trends are observed for the relative bias of $\hat{\gamma}_2$ in Configurations 2 and 3 (second row of Figure 1). The mean relative biases produced by FCS-Standard with either LM or PMM, FCS-LMM-latent, and JM-GL are close to zero. All of the methods show higher variability in relative biases for $\hat{\gamma}_2$ compared to those observed for $\hat{\gamma}_1$ and $\hat{\gamma}_3$. The interquartiles (IQR) of the relative biases of $\hat{\gamma}_2$ in Configurations 2-3 are above 0.20 for all of the imputation methods, whereas the IQRs of the relative bias for $\hat{\gamma}_1$ and $\hat{\gamma}_3$ are around 0.05 and 0.08 on average. This is because BMI is not significantly associated with hospital stay status after adjustment for the other covariates. In Configuration 3, the mean relative biases of $\hat{\gamma}_3$ after using FCS-standard with Poisson regressions, FCS-GLMM, and JM-MGLMM with multilevel Poisson regressions are 0.007, -0.004 and -0.124, respectively. In contrast, the mean relative biases of $\hat{\gamma}_1$ produced by these methods ranges from -0.3 to -0.4.

The relative bias of the subject-level variance, σ_b^2, is depicted in the last row of Figure 1. All of the methods present similar trends across all configurations. The FCS-LMM-latent has the smallest mean relative bias compared to all other methods. The mean relative bias produced by JM-GL is close to zero; however, the variability of the relative bias across replications is the largest. JM-MLMM performs similarly to JM-GL. The two FCS-GLMM methods lead to the largest mean relative biases for estimating σ_b^2 compared to all other methods. Their mean relative biases across all configurations are greater than 0.17.

The relative biases of regression coefficient estimates associated with fully observed predictors are presented in Figure 2. The relative biases of statistically significant coefficients (p-value<0.001) show similar trends for all methods. Generally, the average values and the variability of the relative biases across replications are smaller compared to the relative biases of the statistically insignificant coefficients (the top four rows compared to the bottom two rows). In Configuration 3, the mean relative bias of FCS-Standard is close to zero for all fully observed covariates, except for FCS-Standard with Poisson regression. FCS-LMM-latent has a similar performance to FCS-Standard methods other than the one with Poisson regressions. JM-MGLMM and JM-MLMM have large relative biases for the coefficients of gender, an individual's status of taking prescribed medicines, and having medical bills being paid off over time.

5.3.2 | The RMSE, Interval Width and Coverage

Because all of the methods have produced similar trends across all configurations, we only present the RMSEs, Interval widths, and 95% coverage rates for all imputations methods in Configuration 3 (Table 5). Results for Configurations 1 and 2 are provided in the Appendix. Generally, FCS-LMM-latent, FCS-Standard with either LM or PMM, and JM-GL outperform the other methods in terms of RMSE and average interval width across all regression coefficients. FCS-LMM-latent has the smallest RMSE when

estimating γ_1, γ_2, γ_7, γ_8, γ_9, and smaller RMSEs when estimating other coefficients compared to FCS-Standard with Poisson regression, FCS-GLMM, JM-MLMM, and JM-MGLMM. Across all methods, the average interval widths for all coefficients are similar, and the differences in RMSE are mainly driven by biases.

The interval coverages of all coefficients estimated using FCS-Standard methods, FCS-LMM-latent, and JM-GL are above nominal. The two JM-MLMM methods produce comparable RMSE, average interval widths, and coverages to FCS-LMM-latent only for γ_2 and γ_4. However, JM-MLMM with heteroscedastic Σ_{ei} has significantly lower than nominal coverage for γ_1 and γ_3. Compared to other methods, FCS-GLMM and JM-MGLMM with multilevel Poisson regressions have shorter interval widths for all coefficients and comparable RMSE for most coefficients. However, they result in large RMSEs and below nominal coverages when estimating γ_1 and σ_b^2.

5.3.3 | Results of GLMM for intermittent missing data pattern

In Configuration 3 with intermittent missing data, the FCS-standard methods and JM-GL perform similarly to the one observed for the monotone missing pattern, in terms of the mean relative biases (MRB), interval coverages, and RMSE for most of the coefficients (Appendix table A1). Comparing the monotone to the intermittent missing data pattern, we observe an increase in the RMSE of the coefficients of the incomplete variables for FCS-LMM-latent and FCS-GLMM, and their 95% coverage rates for the co-morbidity index coefficient decrease to 85% (FCS-LMM-latent) and 44% (FCS-GLMM). The performances of the JM-MLMM and JM-MGLMM also deteriorate with intermittent missing data. Specifically, their average 95% coverage rates for the coefficients of the incomplete variables, $\gamma_1 - \gamma_3$, drop from above nominal to below 60%. The two FCS-GLMM methods have higher relative biases, wider interval widths, and interval estimates that do not cover the true parameter for the subject-level variance estimate.

5.4 | Results of the latent growth curve model for Configuration 3

The results of the mean relative biases, RMSE, and average interval coverages of the estimated coefficients associated with the slope factor at each time point are presented in Table 6. Overall, the performances of the JM methods are inferior to those of the FCS methods. Among all methods, JM-MGLMM results in the largest mean relative biases and below nominal interval coverages for all coefficients. Among the FCS methods, the FCS-Standard methods outperform the FCS-LMM-latent and FCS-GLMM. For the comorbidity index coefficient after round 2, FCS-Standard methods result in nominal average interval coverages, while FCS methods with hierarchical modeling result in below nominal coverages. For intermittent missing data patterns, the average interval coverages of FCS-LMM-latent and FCS-GLMM are below 75%. Comparing the results between monotone and intermittent missing data patterns, we observe lower relative biases and better interval coverages for the monotone missing data pattern for all of the FCS methods and JM-GL across all coefficient estimates. JM-MLMM with heteroscedastic variances and

JM-MGLMM result in larger relative biases for the intercepts at rounds 3 and 4 in simulations with intermittent missing data patterns, and their interval estimates do not cover the true intercepts.

5.5 | Results of the multivariate GLMM for Configuration 3

The relative biases of the coefficients estimates of the hospital stay outcome are summarized in Figure 3. For all imputation methods, estimates of model parameters generally have larger relative biases when using a joint analysis model compared to the univariate model analysis. FCS-LMM-latent has the smallest mean relative bias and narrower IQR of relative biases for all coefficients except for the coefficient of the indicator variable of having any prescribed medicine. JM-GL has comparable performance to FCS-LMM-latent in terms of having small mean relative biases, but it has larger variability in relative biases for all coefficients. JM-MLMM and JM-MGLMM lead to small relative biases only in coefficients that are statistically significant (age, BMI, gender, self-rated health). FCS-Standard with either LM or Poisson regressions have higher mean relative biases for coefficients of incomplete variables compared to their performance in the univariate analysis. FCS-Standard with PMM leads to mean relative biases close to zero for coefficients of BMI, gender, self-rated health, and having any prescribed medicine, but it has the widest IQRs of relative biases compared to other imputation methods.

Table 7 summarizes the RMSEs, average interval widths and coverages for the coefficients that represent the associations between incomplete explanatory variables and the two incomplete outcomes. When estimating the conditional association between comorbidity index and the binary outcome hospital stay, λ_{11}, FCS-GLMM with identity links for count variables, FCS-LMM-latent, FCS-GLMM with identity links for count variables, JM-GL, and JM-MLMM have interval coverages above 95%, with small RMSE of approximately 0.04. However, these methods produce larger RMSEs (> 0.11) and lower than nominal interval coverages when estimating the conditional associations between the comorbidity index and the paid assistive devices, λ_{21}. FCS-Standard, FCS-GLMM and JM-MGLMM, which assume Poisson regressions for count variables, lead to small RMSE (0.01, 0.02, 0.01, respectively) and interval coverages around nominal (96%, 92%, 98%, respectively). When estimating the conditional association between BMI and the two outcomes, FCS-LMM-latent produces the smallest RMSE for λ_{21} and λ_{22}, and it has interval coverages above nominal. Generally, all methods lead to higher RMSE when estimating λ_{21} and λ_{22} compared to the estimation of λ_{11} and λ_{12}. This is because of the insignificant association between BMI and the number of paid assistive devices.

For the correlation estimate between the two outcomes, ρ, the FCS-LMM-latent and JM-MLMM have the smallest RMSE compared to the other methods. However, all methods have relatively similar interval estimates which lead to interval coverages that are above nominal.

5.6 | Fraction of missing information and computational time

The FMI estimates of the univariate GLMM analysis are summarized in Table 8. Generally, the average FMI estimates from simulations with monotone missing data patterns are smaller than those with intermittent missing data patterns for all imputation methods. When estimating the coefficients associated with incomplete variables and the subject-level variance, larger differences between the FMI estimated from the intermittent and monotone missing data patterns are observed for FCS-based methods and JM-MLMM with heteroscedastic variances compared to the other imputation methods. Across all simulation scenarios, the range of the average FMI estimates is between 0.14 and 0.51. Most imputation methods have average FMI estimates between 0.20 to 0.30. Among all of the methods, JM-GL results in the smallest FMI estimates for all of the constant effects coefficients.

The average FMI estimates of the coefficients in the latent curve model when assuming a monotone missing data pattern are smaller than those assuming the intermittent missing data pattern (Table 6). Overall, the FMI estimates for the latent growth curve model coefficients are larger than the univariate GLMM coefficients. The FCS-based methods and JM-GL result in relatively smaller FMI that are around 0.25 compared to the JM-MLMM with heteroscedastic variances.

We recorded the computational time of all of the 10 methods on a standard Windows PC intel Xeon Core, 2.40GHz processor. For a single imputation, the FCS-standard with the linear regressions and PMM procedures, and JM-GL required the least computational time (less than 2s). FCS-Standard with the Poisson regressions had an average computational time of 22s. The three JM-based methods with multivariate multilevel modeling required approximately 280s of computation time. The computational time of FCS-GLMM methods was around 800s. The most computationally intensive method was the FCS-LMM-latent, which required several hours to complete a single imputation.

6 | CASE STUDY ON SKILLED NURSING HOME ADMISSIONS

We applied the different imputation methods to study the relationship between sociodemographic factors, the number of hospital admissions, measures of physical, cognitive abilities, and skilled nursing facilities admissions. The data are based on NHATS Rounds 1-5, linked to data from the Center for Medicare and Medicaid (CMS). We selected 4836 participants who were enrolled in the Medicare fee for service at baseline. Deceased participants and those who left the Medicare fee for service during the study were assumed to be censored in contrast to missing. The proportions of missing participants because of loss to follow up at each round were approximately 16%, 14%, 12%, and 6%.

The sociodemographic factors comprised baseline age, sex, race/ethnicity (non-Hispanic white and other), educational level (college and higher, no school or < 9 grade, 9-12 grade and high school, and vocational training), and time-varying cohabitation status (living alone versus not living alone) and Medicare-Medicaid dual eligibility. The cognitive ability is a 5-points Likert scale ranging from excellent (1 point) to poor (5 points). The status of time-varying Alzheimer's Disease and Related Dementias

(ADRD) was defined based on the Chronic Condition Data Warehouse (CCW) algorithm. The time-varying comorbidity index comprises a count of 20 common chronic conditions among older adults identified through the CCW algorithm. [64] The function and frailty factors include two time-varying count variables: activities of daily living (ADL) and frailty score. ADL is derived as a count of limitations in seven daily living activities. Frailty score is defined by the sum of five criteria: exhaustion, low activity, weakness, slowness, and shrinking. Among all explanatory variables, only cohabitation status, ADL, and frailty score are subject to missing values.

We implemented FCS-Standard, FCS-LMM-latent, FCS-GLMM, JM-GL, JM-MLMM, and JM-MGLMM to impute the missing data. We applied a multilevel logistic regression model to examine the relationship between SNF admission status and sociodemographic factors adjusted for other factors that assess the physical and cognitive abilities. The results are displayed in Figure 4. Patients who are non-Hispanic whites are associated with 75% higher conditional odds of having an SNF admission compared to other race/ethnicity groups. For one unit increase in the number of hospital admissions, the conditional odds of having an SNF admission will quadruple. Being diagnosed with ADRD, with a higher ADL disabilities value and a higher comorbidity index, especially frailty, were associated with increased conditional odds of SNF admission. Not living alone and having Medicare-Medicaid dual eligibility were associated with decreased conditional odds of SNF. Education showed a statistically insignificant relationship with any SNF admission after adjustment for the other covariates.

All imputation methods resulted in similar estimates of the conditional odds ratio and 95% interval estimates for the fully observed variables. FCS-based methods and JM-based methods result in different estimates for variables with missing values. FCS-Standard produces higher conditional odds ratios for cohabitation status and the number of hospital admissions, and it has lower conditional odds ratios for ADL disabilities and frailty scores compared to other methods. Compared to FCS-based methods, JM-based methods have higher conditional odds ratios for estimates of ADL disabilities and frailty scores.

7 | DISCUSSION

Addressing missing data in longitudinal studies is challenging and involves advanced statistical techniques. We review existing MI methods for incomplete longitudinal mixed data, and their implementations on widely accessible software that requires limited additional coding. We compare these methods using simulation analyses and describe an applied example based on the NHATS data.

Among all of the methods examined, the FCS-LMM-latent method had the best performance for the univariate and multivariate multilevel modeling in terms of having small biases, relatively low FMI, and interval coverages that were at or above nominal. FCS-Standard and JM-GL resulted in the best performance for the growth curve model analysis, and comparable performance to FCS-LMM-latent for univariate hierarchical regression modeling, but sub-optimal performance when estimating some of the

coefficients of multivariate multilevel modeling. Across all analyses, all imputation methods displayed better performance for data with monotone missing data patterns compared to intermittent missing data. This is partly because our intermittent missing data pattern is NMAR, while the monotone missing data pattern examined is MAR.

FCS-standard and JM-GL posit that the same variable measured at a different time point is a different variable, and use conditional associations between variables measured at different time points to impute the ones that are incomplete. This process essentially assumes an unstructured association structure between the repeated measurements. FCS-LMM-latent models adjust for the associations between repeated measurements recorded at different time points through subject-specific intercepts. The FCS-Standard methods are computationally more efficient and more flexible for imputing non-Normally distributed variables compared to FCS methods with multilevel modeling. The disadvantage of FCS-Standard methods is that they can result in non-identifiable models as the number of waves and variables within waves increases. In such cases, the FCS-twofold method[41] is suggested for building imputation models. In our simulation study, we did not implement the FCS-twofold method because the simulation data comprise a small number of waves and variables within each wave. Possible disadvantages of the FCS-LMM-latent method are its intensive computational time and its possible sub-optimal performance when estimating growth curve models.

The JM-based methods with multilevel modeling, JM-MLMM-latent, and JM-MGLMM use fully observed variables as predictors and model the associations between incomplete variables through the correlations of the subject-level intercepts. This may result in biased estimates of the association between an outcome and a predictor in the analysis model when both variables are incomplete. In addition, the simulations display that using JM-MLMM-latent with a common covariance Σ_e across subjects had slightly better performance compared to JM-MLMM-latent with heterogeneous-covariances Σ_{e_i}. The poor performance of JM-MLMM-latent with heterogeneous-covariances may arise from the lack of convergence of the sampling algorithm. The Gibbs sampler used by JM-MLMM-latent (random) is slow to converge with a large number of subjects and poorly estimated subject-level variance.[14]

Imputation of count variables is sensitive to distributional assumptions. The simulations show that methods using Poisson regression may produce biased estimates of the comorbidity coefficient compared to a multilevel linear model or predictive mean matching. This may arise because many subjects have no comorbidities and the total number of comorbidities is bounded. Thus, a Poisson or a Negative-Binomial regression does not approximate this variable well. In situations where Poisson or Negative-Binomial hierarchical regression models approximate the data well (e.g. the number of assistive devices), our simulations show that these methods perform similarly to multilevel linear regression or predictive mean matching. Another option for imputing non-Normal continuous variables is to transform them so they would appear approximately Normal during the imputation step and transform the variables back in the statistical analyses step. However, these transformations may not preserve the original correlation structure between variables.[19,65]

In univariate analysis, FCS-GLMM methods had point estimates of regression coefficients with small biases, but it resulted in poor operating characteristics for point and interval estimates of subject-level variance when using a multilevel model for analysis. This may arise from the computational algorithm that is being used to impute binary variables. The MCMC algorithm that has been used to sample from a multilevel logistic regression in the `micemd` package samples $\mathbf{b}_i^*$ from its marginal posterior distribution $N(0, \boldsymbol{\Psi}_b)$. This can result in underestimation of the variance of $\mathbf{b}_i$ in the imputation process which carries over to the analysis model.

In our simulations, FCS-Standard performs well in cases with missing predictors but results in worse operating characteristics if these incomplete variables are used as dependent variables. This finding relates to the differences between the imputation of explanatory variables and outcomes. When both explanatory variables and outcomes are missing, Little[66] suggested that imputing outcome variables provides limited information for the subsequent regression of analysis. Von Hippel[67] proposed the multiple-imputation-deletion procedure that includes all incomplete variables in the imputation step and excludes missing response variables values from the substantive analysis. However, this procedure was proposed when the response variables and the subsequent analyses are defined in advance. In this work, we considered situations where no specific statistical analysis is specified before the imputation step, and the imputed datasets can be used for multiple analyses.

Our results agree with the conclusions of Huque et al[21] that FCS-Standard provides reliable estimates for univariate hierarchical model analysis. However, we also demonstrated that FCS-Standard methods produce higher bias for coefficient estimates of level-1 covariates when analyzing multiple outcomes simultaneously in joint multilevel models. When imputing count variables, our results are consistent with the findings of Demirtas et al[65] and Kalaycioglu et al,[19] who demonstrated that assuming a linear regression model for a non-Normally distributed variable may result in smaller biases compared to a mis-specified non-Gaussian model. Our simulations show that imputations using either the predictive mean matching or a multilevel linear model with rounding result in good operating characteristics of subsequent analyses.

In addition to the methods discussed in the manuscript, there are methods that can be used for imputing incomplete longitudinal data, but require more technical coding and advanced statistical knowledge. One option is the fully Bayesian approach.[19] It requires users to specify prior on the parameters of both imputation and analysis models and implement additional coding with software, such as `STAN` or `PROC MCMC` in `SAS`. The MCMC algorithm updates the parameters of the imputation model, draws an imputation for the incomplete variable, and updates the parameters of the analysis model sequentially at each iteration. More details on this approach can be found in Kalaycioglu et al.[19] Another option is to implement nonparametric modeling techniques, such as sequential regression trees, random forests, and Dirichlet process mixture models.[68–71] These methods are beyond the scope of this manuscript, because their implementation with multiple variables that experience missing values is not trivial and

they require the use of advanced statistical theory and coding. A future direction of our comparative work is to consider non-parametric imputation methods and extend the study simulations to data with more incomplete variables measured at a larger number of waves.

In conclusion, in longitudinal data with a small number of waves and a limited number of variables, when the analysis models comprise univariate regression models, FCS-standard is a computationally efficient method that results in precise and accurate estimates for both single and multilevel models. However, if the analysis models comprise multivariate multilevel models FCS-LMM-latent is a valid statistical method that produces more accurate estimates at the cost of more intensive computations.

ACKNOWLEDGMENTS

DATA AVAILABILITY STATEMENT

The data that support the findings of the simulation study are openly available from the National Health and Aging Trends Study (NHATS) at https://nhats.org/researcher/data-access/public-use-files. The data that support the findings of the real data analysis are available at https://nhats.org/researcher/data-access/sensitive-data-files?id=restricted_data_files. Restrictions apply to the availability of these data, which were used under license for this study.

FINANCIAL DISCLOSURE

None reported.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

References

1. Molenberghs G, Fitzmaurice G, Kenward MG, Tsiatis A, Verbeke G. *Handbook of Missing Data Methodology*. Chapman and Hall/CRC . 2014.

2. Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 1977; 39(1): 1-22.

3. Tanner MA, Wong WH. The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association* 1987; 82(398): 528–540.

4. Little RJA. Missing-Data Adjustments in Large Surveys. *Journal of Business & Economic Statistics* 1988; 6(3): 287–296.

5. Holt D, Elliot D. Methods of Weighting for Unit Non-Response. *Journal of the Royal Statistical Society. Series D (The Statistician)* 1991; 40(3): 333–342.

6. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Statistical Methods in Medical Research* 2013; 22(3): 278-295.

7. Little RJA, Rubin DB. *Statistical Analysis with Missing Data, Third Edition*. USA: John Wiley Sons, Inc. . 2019.

8. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons . 1987.

9. Rubin DB. Inference and Missing Data. *Biometrika* 1976; 63(3): 581–592.

10. Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology* 2001; 27.

11. Schafer J. *Analysis of incomplete multivariate data*. Chapman&Hall . 1997.

12. Schafer JL, Olsen MK. Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective. *Multivariate Behavioral Research* 1998; 33(4): 545-571.

13. Van Buuren S, Brand JP, Groothuis-Oudshoorn C, Rubin DB. Fully conditional specification in multivariate imputation. *J Stat Comput Simul* 2006; 76.

14. Yucel RM. Random-covariances and mixed-effects models for imputing multivariate multilevel continuous data. *Statistical modelling* 2001; 11(4).

15. Yucel RM. Multiple imputation inference for multivariate multilevel continuous data with ignorable non-response. *Philosophical transactions Series A, Mathematical, physical, and engineering sciences* 2008; 366(1874): 2389-403.

16. Yucel RM, Zhao E, Schenker N, Raghunathan TE. Sequential Hierarchical Regression Imputation. *Journal of Survey Statistics and Methodology* 2017; 6(1): 1-22.

17. Schafer JL, Yucel RM. Computational Strategies for Multivariate Linear Mixed-Effects Models With Missing Values. *Journal of Computational and Graphical Statistics* 2002; 11(2): 437-457.

18. Jolani S. Hierarchical imputation of systematically and sporadically missing data: An approximate Bayesian approach using chained equations. *Biometrical Journal* 2018; 60(2): 333-351.

19. Kalaycioglu O, Copas A, King M, Omar RZ. A comparison of multiple-imputation methods for handling missing data in repeated measurements observational studies. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2016; 179(3): 683-706.

20. Buuren S, Oudshoorn C. Multivariate Imputation by Chained Equations : Mice V1.0 User's manual. In: ; 2000.

21. Huque MH, Carlin JB, Simpson JA, Lee KJ. A comparison of multiple imputation methods for missing data in longitudinal studies. *BMC Medical Research Methodology* 2018; 18(1): 168.

22. Jolani S, Debray TPA, Koffijberg H, Van Buuren S, Moons KGM. Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. *Statistics in Medicine* 2015; 34(11): 1841-1863.

23. Audigier V, White IR, Jolani S, Debray TP, Quartagno M, Carpenter J. Multiple imputation for multilevel data with continuous and binary variables. *Stat Sci* 2018; 33.

24. Zhao E, Yucel RM. Performance of sequential imputation method in multilevel applications. *Proceedings in Joint statistical meetings Washington DC. 2009.* 2009.

25. Resche-Rigon M, White IR. Multiple imputation by chained equations for systematically and sporadically missing multilevel data. *Statistical methods in medical research* 2018; 27(6).

26. Enders CK, Mistler SA, Keller BT. Multilevel multiple imputation: a review and evaluation of joint modeling and chained equations imputation. *Psychol Methods* 2016; 21.

27. Kasper JD, Freedman VA. Findings From the 1st Round of the National Health and Aging Trends Study (NHATS): Introduction to a Special Issue. *The Journals of Gerontology: Series B* 2014; 69(Suppl_1): S1-S7.

28. R Core Team . *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; Vienna, Austria: 2020.

29. Cao Y. Github repository: Multiple imputation for longitudinal data. 2021. https://github.com/Yi-Cao1227/Multiple-Imputation-for-longitudinal-data.

30. Carpenter JR, Kenward MG, White IR. Sensitivity analysis after multiple imputation under missing at random: a weighting approach. *Statistical Methods in Medical Research* 2007; 16(3): 259-275.

31. Andridge RR. Quantifying the impact of fixed effects modeling of clusters in multiple imputation for cluster randomized trials. *Biometrical Journal* 2011; 53(1): 57-74.

32. Drechsler J. Multiple Imputation of Multilevel Missing Data—Rigor Versus Simplicity. *Journal of Educational and Behavioral Statistics* 2015; 40(1): 69-95.

33. Grund S, Lüdtke O, Robitzsch A. Multiple Imputation of Missing Data for Multilevel Models: Simulations and Recommendations. *Organizational Research Methods* 2018; 21(1): 111-149.

34. Gelman A. Analysis of variance—why it is more important than ever. *The annals of statistics* 2005; 33(1): 1–53.

35. Buuren SV. *Flexible Imputation of Missing Data, Second Edition*. Chapman and Hall/CRC . 2018.

36. Buuren SV, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, Articles* 2011; 45(3): 1–67.

37. Raghunathan TE, Hoewyk JV. IVEware: Imputation and Variance Estimation Software User Guide. 2002.

38. Andridge RR, Little RJA. A Review of Hot Deck Imputation for Survey Non-response. *International Statistical Review* 2010; 78(1): 40-64.

39. Kleinke K, Reinecke J. Multiple imputation of incomplete zero-inflated count data. *Statistica Neerlandica* 2013; 67(3): 311-336.

40. Kleinke K, Reinecke J. countimp 1.0 – A Multiple Imputation Package for Incomplete Count Data. 2013.

41. Nevalainen J, Kenward MG, Virtanen SM. Missing values in longitudinal dietary data: A multiple imputation approach based on a fully conditional specification. *Statistics in Medicine* 2009; 28(29): 3657-3669.

42. Welch CA, Petersen I, Bartlett JW, et al. Evaluation of two-fold fully conditional specification multiple imputation for longitudinal electronic health record data. *Statistics in Medicine* 2014; 33(21): 3725-3737.

43. Fong Y, Rue H, Wakefield J. Bayesian inference for generalized linear mixed models. *Biostatistics* 2010; 11(3): 397-412.

44. Horton NJ, Lipsitz SR, Parzen M. A Potential for Bias When Rounding in Multiple Imputation. *The American Statistician* 2003; 57(4): 229-232.

45. Yucel RM, He Y, Zaslavsky AM. Using Calibration to Improve Rounding in Imputation. *The American Statistician* 2008; 62(2): 125-129. doi: 10.1198/000313008X300912

46. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. Chapman and Hall/CRC. 3rd ed. 2013.

47. Goldstein H, Carpenter J, Kenward MG, Levin KA. Multilevel models with multivariate mixed response types. *Stat Model* 2009; 9.

48. Enders CK, Keller BT, Levy R. A fully conditional specification approach to multilevel imputation of categorical and continuous variables. *Psychological methods* 2018; 23(2).

49. Aitchison J, Bennett JA. Polychotomous quantal response by maximum indicant. *Biometrika* 1970; 57(2): 253-262.

50. Keller BT, Enders CK. *Blimp Software Manual (Version Beta 6.7). Los Angeles, Ca* . 2017.

51. Belin TR, Hu MY, Young AS, Grusky O. Performance of a general location model with an ignorable missing-data assumption in a multivariate mental health services study. *Statistics in Medicine* 1999; 18(22): 3123-3135.

52. Schafer JL. mix: Estimation/Multiple Imputation for Mixed Categorical and Continuous Data. 2017. R package version 1.0-10.

53. Zhao JH, Schafer JL. *pan: Multiple imputation for multivariate panel or clustered data*. 2018. R package version 1.6.

54. Carpenter JR, Kenward MG. *Multiple Imputation and its Application, First Edition*. John Wiley & Sons, Ltd . 2013.

55. Quartagno M, Carpenter J. *jomo: A package for Multilevel Joint Modelling Multiple Imputation*. 2014.

56. Carpenter JR, Goldstein H, Kenward MG. REALCOM-IMPUTE software for multilevel multiple imputation with mixed response types. *J Stat Softw* 2011; 45.

57. Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. 2003.

58. Duncan TE, Duncan SC. An introduction to latent growth curve modeling. *Behavior Therapy* 2004; 35(2): 333-363. doi: https://doi.org/10.1016/S0005-7894(04)80042-X

59. Verbeke G, Fieuws S, Molenberghs G, Davidian M. The analysis of multivariate longitudinal data: a review.. *Statistical methods in medical research* 2014; 23(1): 42–59.

60. Jaffa MA, Gebregziabher M, Jaffa AA. Analysis of multivariate longitudinal kidney function outcomes using generalized linear mixed models. *Journal of Translational Medicine* 2015; 13(1): 192.

61. Cao Y, Allore H, Gutman R, Vander Wyk B, Jørgensen TSH. Risk Factors of Skilled Nursing Facility Admissions and the Interrelation With Hospitalization and Amount of Informal Caregiving Received. *Medical Care* 2022; 60(4).

62. Meng XL. Multiple-imputation inferences with uncongenial sources of input. *Statistical Science* 1994: 538–558.

63. Larose C, Dey DK, Harel O. THE IMPACT OF MISSING VALUES ON DIFFERENT MEASURES OF UNCERTAINTY. *Statistica Sinica* 2019; 29(2): 551–566.

64. Goodman RA, Posner SF, Huang ES, Parekh AK, Koh HK. Defining and Measuring Chronic Conditions: Imperatives for Research, Policy, Program, and Practice. *Prev Chronic Dis* 2013; 10: E66.

65. Demirtas H, Freels SA, Yucel RM. Plausibility of multivariate normality assumption when multiply imputing non-Gaussian continuous outcomes: a simulation assessment. *Journal of Statistical Computation and Simulation* 2008; 78(1): 69-84.

66. Little RJA. Regression With Missing X's: A Review. *Journal of the American Statistical Association* 1992; 87(420): 1227–1237.

67. Von Hippel PT. REGRESSION WITH MISSING YS: AN IMPROVED STRATEGY FOR ANALYZING MULTIPLY IMPUTED DATA. *Sociological Methodology* 2007; 37(1): 83-117.

68. Burgette LF, Reiter JP. Multiple Imputation for Missing Data via Sequential Regression Trees. *American Journal of Epidemiology* 2010; 172(9): 1070-1076. doi: 10.1093/aje/kwq260

69. Vidotto D, Vermunt JK, Van Deun K. Bayesian Latent Class Models for the Multiple Imputation of Categorical Data. *Methodology* 2018; 14(2): 56-68. doi: 10.1027/1614-2241/a000146

70. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2011; 28(1): 112-118. doi: 10.1093/bioinformatics/btr597

71. Wongkamthong C, Akande O. A Comparative Study of Imputation Methods for Multivariate Ordinal Data. *Journal of Survey Statistics and Methodology* 2021. doi: 10.1093/jssam/smab028

ID	$\mathbf{Y}_1$	$\mathbf{Y}_2$	$\mathbf{Y}_3$	Time
1	y_{111}	y_{121}	y_{131}	t_1
1	y_{112}	y_{122}	y_{132}	t_2
1	?	?	?	t_3
1	?	?	?	t_4
2	y_{211}	y_{221}	y_{231}	t_1
2	y_{212}	y_{222}	y_{232}	t_2
2	y_{213}	y_{223}	y_{233}	t_3
2	?	?	?	t_4
3	y_{311}	y_{321}	y_{331}	t_1
3	y_{312}	y_{322}	y_{332}	t_2
3	?	y_{323}	?	t_3
3	y_{314}	y_{324}	y_{334}	t_4

TABLE 2 Example of data arranged in the long format

ID	Time t_1			Time t_2			Time t_3			Time t_4		
	$\mathbf{Y}_{11}$	$\mathbf{Y}_{21}$	$\mathbf{Y}_{31}$	$\mathbf{Y}_{12}$	$\mathbf{Y}_{22}$	$\mathbf{Y}_{32}$	$\mathbf{Y}_{13}$	$\mathbf{Y}_{23}$	$\mathbf{Y}_{33}$	$\mathbf{Y}_{14}$	$\mathbf{Y}_{24}$	$\mathbf{Y}_{34}$
1	y_{111}	y_{121}	y_{131}	y_{112}	y_{122}	y_{132}	?	?	?	?	?	?
2	y_{211}	y_{221}	y_{231}	y_{212}	y_{222}	y_{232}	y_{213}	y_{223}	y_{233}	?	?	?
3	y_{311}	y_{321}	y_{331}	y_{312}	y_{322}	y_{332}	?	y_{323}	?	y_{314}	y_{324}	y_{334}

TABLE 1 Example of data arranged in the wide format

Approach	Data format	Method	Imputation models	
FCS			Imputation model for binary variables	Imputation model for count variables
	Wide	FCS-Standard (LM)	Logistic regression	Linear regression
		FCS-Standard (PMM)	Logistic regression	Predictive mean matching
		FCS-Standard (Poisson)	Logistic regression	Poisson regression
	Long	FCS-LMM-latent	Multilevel linear regression on latent variables	Multilevel linear regression
		FCS-GLMM (Gaussian)	Multilevel logistic regression	Multilevel linear regression
		FCS-GLMM (Poisson)	Multilevel logistic regression	Multilevel Poisson regression
JM	Wide	JM-GL	General location model	
	Long	JM-MLMM-latent (common)	Multivariate multilevel linear model with latent variables and homoscedastic within-subject variance	
		JM-MLMM-latent (random)	Multivariate multilevel linear model with latent variables and heteroscedastic within-subject variance	
		JM-MGLMM	Multivariate multilevel generalized linear model using a logit and log link for binary, count variables	

TABLE 3 Summary of imputation methods for mixed-type longitudinal data

Configuration	Incomplete variables	Missing Pattern	Statistical Analysis
1	Hospital stay (Binary) Comorbidity index (Count)	Monotone	Univariate GLMM
2	Hospital stay (Binary) Comorbidity index (Count) BMI (Continuous)	Monotone	Univariate GLMM
3	Hospital stay (Binary) Comorbidity index (Count) BMI (Continuous) Number of devices paid for caring (Count)	Monotone Intermittent	Univariate GLMM Growth Curve Model Bivariate GLMM

TABLE 4 Summary of simulation design

Estimates	Metrics	FCS-Standard (LM)	FCS-Standard (PMM)	FCS-Standard (Poisson)	FCS-LMM-latent	FCS-GLMM (Gauss.)	FCS-GLMM (Poisson)	JM-General Location	JM-MLMM (common)	JM-MLMM (random)	JM-MGLMM
Intercept $\hat{\gamma}_0^{comp} = -3.89^{**}$	RMSE	0.150	0.154	0.149	0.156	0.255	0.237	0.176	0.131	0.165	0.125
	Width	1.157	1.136	1.127	1.110	1.039	1.033	1.094	1.131	1.144	1.055
	Coverage(%)	100	99	100	100	99	100	100	100	100	100
Comorbidity Index $\hat{\gamma}_1^{comp} = 0.38^{**}$	RMSE	0.021	0.020	0.136	0.019	0.025	0.112	0.020	0.031	0.071	0.152
	Width	0.117	0.118	0.117	0.117	0.110	0.105	0.113	0.117	0.113	0.100
	Coverage(%)	99	100	0	100	100	0	99	100	22	0
BMI $\hat{\gamma}_2^{comp} = -0.02^{*}$	RMSE	0.003	0.004	0.004	0.003	0.004	0.003	0.004	0.004	0.006	0.005
	Width	0.029	0.029	0.028	0.028	0.026	0.026	0.027	0.028	0.030	0.027
	Coverage(%)	100	100	100	100	100	100	100	100	99	100
Paid assistive devices $\hat{\gamma}_3^{comp} = 0.36^{**}$	RMSE	0.026	0.022	0.021	0.026	0.019	0.018	0.025	0.019	0.052	0.048
	Width	0.149	0.152	0.150	0.150	0.142	0.141	0.144	0.149	0.151	0.143
	Coverage(%)	100	100	100	99	100	100	99	100	87	92
Age $\hat{\gamma}_4^{comp} = 0.11^{**}$	RMSE	0.014	0.014	0.015	0.016	0.018	0.015	0.019	0.017	0.018	0.012
	Width	0.112	0.112	0.111	0.111	0.104	0.101	0.107	0.112	0.111	0.106
	Coverage(%)	100	100	100	100	100	100	100	100	100	100
Gender $\hat{\gamma}_5^{comp} = 0.28^{**}$	RMSE	0.044	0.041	0.073	0.056	0.048	0.072	0.053	0.040	0.054	0.090
	Width	0.310	0.311	0.306	0.307	0.282	0.282	0.299	0.307	0.304	0.287
	Coverage(%)	100	100	99	100	100	96	100	100	100	96
Self-rated health $\hat{\gamma}_6^{comp} = 0.20^{**}$	RMSE	0.024	0.024	0.052	0.024	0.021	0.049	0.028	0.022	0.040	0.072
	Width	0.162	0.162	0.164	0.157	0.146	0.146	0.155	0.157	0.158	0.148
	Coverage(%)	100	99	95	100	100	93	100	100	98	60
Have med. bill paid off overtime $\hat{\gamma}_7^{comp} = 0.34^{*}$	RMSE	0.064	0.066	0.060	0.048	0.058	0.066	0.131	0.093	0.102	0.067
	Width	0.483	0.483	0.480	0.473	0.444	0.435	0.462	0.473	0.475	0.451
	Coverage(%)	100	100	100	100	100	100	93	99	99	100
Took prescribed med. $\hat{\gamma}_8^{comp} = 0.30$	RMSE	0.095	0.096	0.129	0.071	0.096	0.156	0.110	0.155	0.206	0.183
	Width	0.702	0.714	0.692	0.666	0.647	0.644	0.667	0.684	0.674	0.650
	Coverage(%)	100	100	100	100	100	100	99	100	99	98
Time $\hat{\gamma}_9^{comp} = 0.10^{**}$	RMSE	0.020	0.022	0.048	0.016	0.027	0.023	0.021	0.016	0.017	0.018
	Width	0.109	0.108	0.108	0.104	0.102	0.105	0.099	0.105	0.106	0.100
	Coverage(%)	99	100	66	100	96	98	99	100	100	100
Random effect $\hat{\sigma}_{b_{comp}}^2 = 1.13^{**}$	RMSE	0.048	0.044	0.070	0.037	0.202	0.220	0.050	0.035	0.040	0.107
	Width	0.235	0.239	0.233	0.238	0.217	0.212	0.221	0.241	0.240	0.220
	Coverage(%)	98	99	95	100	0	0	94	100	99	56

TABLE 5 The root mean squared error (RMSE), average interval width and empirical coverage of the 95% CI of coefficients estimates and subject-level variance estimate for the univariate analysis in Configuration 3 with the monotone missing data pattern. The incomplete predictors are marked in bold.

Slope Estimates	Metrics	FCS-Standard (LM)		FCS-Standard (PMM)		FCS-Standard (Poisson)		FCS-LMM-latent		FCS-GLMM (Gaussian)		FCS-GLMM (Poisson)		JM-General Location		JM-MLMM (common)		JM-MLMM (random)		JM-MGLMM	
		Inter.	Mono.	Inter.	Mono.	Inter.	Mono.	Inter.	Mono.	Inter.	Mono.	Inter.	Mono.	Inter.	Mono.	Inter.	Mono.	Inter.	Mono.	Inter.	Mono.
Round 1 Intercept	MRB	-0.03	0.00	-0.02	0.00	-0.03	-0.01	0.11	0.06	0.11	0.06	0.11	0.06	-0.03	-0.01	0.11	0.05	0.16	0.32	0.89	0.05
	RMSE	0.21	0.16	0.21	0.17	0.21	0.16	0.22	0.15	0.22	0.15	0.22	0.15	0.21	0.16	0.22	0.15	0.31	0.56	1.63	0.15
	Coverage(%)	74	85	77	87	72	86	60	85	60	84	58	88	71	83	61	88	64	26	13	93
	FMI	0.40	0.32	0.47	0.46	0.38	0.31	0.31	0.24	0.29	0.26	0.29	0.25	0.26	0.23	0.30	0.25	0.44	0.68	0.08	0.34
Round 1 Comorbidity	MRB	-0.43	0.08	-0.48	0.40	-2.15	1.96	-4.01	-1.57	-4.14	-1.54	-4.07	-2.87	-0.48	-0.15	-1.19	-0.05	-1.57	4.84	-12.53	-2.47
	RMSE	0.03	0.02	0.03	0.02	0.04	0.04	0.06	0.03	0.06	0.03	0.06	0.05	0.03	0.02	0.03	0.02	0.11	0.20	0.19	0.04
	Coverage(%)	100	100	100	100	94	94	70	98	65	98	44	82	97	100	100	100	93	97	0	96
	FMI	0.42	0.31	0.42	0.31	0.45	0.53	0.37	0.26	0.35	0.28	0.38	0.50	0.25	0.23	0.38	0.28	0.84	0.92	0.14	0.41
Round 2 Intercept	MRB	0.01	0.01	0.02	0.00	0.02	0.01	0.00	0.01	0.00	0.01	0.00	0.02	0.01	0.01	0.00	0.02	0.46	0.37	4.75	0.23
	RMSE	0.18	0.16	0.19	0.16	0.18	0.16	0.12	0.12	0.12	0.12	0.12	0.12	0.18	0.16	0.12	0.13	0.94	0.76	9.50	0.48
	Coverage(%)	64	71	84	81	66	71	83	79	85	79	85	80	60	66	83	78	09	11	0	5
	FMI	0.42	0.36	0.74	0.59	0.43	0.34	0.31	0.28	0.34	0.28	0.32	0.29	0.30	0.27	0.34	0.29	0.93	0.89	0.24	0.45
Round 2 Comorbidity	MRB	-0.23	-0.08	-0.24	0.01	-0.86	0.41	-1.34	-0.61	-1.37	-0.59	-1.39	-1.06	-0.23	-0.12	-0.45	-0.30	-0.36	1.37	-2.31	-0.93
	RMSE	0.03	0.02	0.04	0.02	0.05	0.03	0.08	0.04	0.08	0.04	0.08	0.06	0.04	0.02	0.04	0.03	0.13	0.23	0.14	0.06
	Coverage(%)	99	100	98	100	77	95	39	94	38	96	9	62	91	100	99	99	89	96	9	63
	FMI	0.52	0.33	0.53	0.33	0.60	0.64	0.44	0.32	0.44	0.34	0.52	0.64	0.32	0.25	0.47	0.34	0.90	0.95	0.24	0.51
Round 3 Intercept	MRB	-0.04	-0.01	-0.02	-0.03	-0.04	-0.01	0.17	0.10	0.17	0.10	0.17	0.10	-0.04	-0.01	0.17	0.10	1.09	0.74	9.47	0.52
	RMSE	0.20	0.14	0.20	0.14	0.19	0.14	0.30	0.20	0.31	0.20	0.30	0.21	0.20	0.14	0.30	0.20	1.85	1.26	16.04	0.89
	Coverage(%)	62	76	76	83	62	75	13	43	14	44	19	43	52	70	18	45	0	0	0	0
	FMI	0.55	0.39	0.77	0.62	0.53	0.39	0.39	0.35	0.40	0.34	0.41	0.34	0.36	0.29	0.41	0.36	0.97	0.94	0.22	0.50
Round 3 Comorbidity	MRB	-0.21	-0.11	-0.22	-0.05	-0.66	0.13	-1.01	-0.50	-1.03	-0.48	-1.07	-0.79	-0.21	-0.13	-0.44	-0.36	-0.31	0.76	-0.99	-0.73
	RMSE	0.04	0.02	0.04	0.02	0.07	0.03	0.10	0.05	0.10	0.05	0.10	0.08	0.05	0.03	0.05	0.04	0.16	0.24	0.10	0.07
	Coverage(%)	95	99	97	99	54	97	16	81	17	82	1	39	82	99	90	92	84	93	46	25
	FMI	0.59	0.37	0.60	0.37	0.71	0.73	0.49	0.35	0.49	0.37	0.60	0.70	0.36	0.29	0.52	0.37	0.92	0.96	0.31	0.54
Round 4 Intercept	MRB	-0.03	-0.02	-0.04	-0.03	-0.02	-0.01	0.10	0.02	0.10	0.02	0.10	0.02	-0.03	-0.02	0.09	0.02	1.04	0.60	9.99	0.45
	RMSE	0.33	0.21	0.33	0.21	0.33	0.21	0.24	0.14	0.24	0.14	0.24	0.14	0.34	0.22	0.23	0.14	2.03	1.19	19.36	0.88
	Coverage(%)	62	78	70	87	63	79	72	92	72	91	69	92	54	70	73	91	1	5	0	0
	FMI	0.61	0.40	0.76	0.70	0.61	0.43	0.42	0.32	0.42	0.32	0.41	0.33	0.40	0.30	0.43	0.32	0.95	0.88	0.16	0.45
Round 4 Comorbidity	MRB	-0.32	-0.16	-0.33	-0.10	-0.81	0.02	-1.06	-0.50	-1.09	-0.47	-1.14	-0.82	-0.30	-0.17	-0.48	-0.42	-0.29	0.80	-1.18	-0.76
	RMSE	0.05	0.03	0.06	0.03	0.08	0.04	0.10	0.05	0.11	0.05	0.11	0.08	0.06	0.03	0.05	0.05	0.18	0.26	0.12	0.07
	Coverage(%)	93	99	94	99	58	97	23	87	22	88	3	46	76	97	92	93	84	93	50	35
	FMI	0.61	0.39	0.63	0.39	0.74	0.74	0.49	0.34	0.50	0.36	0.59	0.65	0.36	0.30	0.53	0.35	0.91	0.95	0.35	0.52

TABLE 6 The mean relative bias (MRB), the root mean squared error (RMSE), average empirical coverage of the 95% CI, average FMI estimates of slope factor's coefficients estimates for the latent growth curve analysis in Configuration 3 with intermittent (Inter.) and monotone (Mono.) missing data patterns.

Estimates	Metrics	FCS-Standard (LM)	FCS-Standard (PMM)	FCS-Standard (Poisson)	FCS-LMM-latent	FCS-GLMM (Gauss.)	FCS-GLMM (Poisson)	JM-General Location	JM-MLMM (common)	JM-MLMM (random)	JM-MGLMM
					Outcome 1: Hospital stay (Yes)						
Comorbidity $\hat{\lambda}_{11}^{comp} = 0.44^{**}$	RMSE	0.064	0.160	0.143	0.043	0.038	0.113	0.070	0.046	0.059	0.193
	Width	0.216	0.890	0.520	0.482	0.768	0.614	0.596	1.140	0.633	0.339
	Coverage	81	88	85	100	100	80	100	100	100	35
BMI $\hat{\lambda}_{12}^{comp} = -0.02^{**}$	RMSE	0.006	0.008	0.006	0.004	0.005	0.005	0.006	0.005	0.007	0.007
	Width	0.032	0.062	0.049	0.042	0.055	0.052	0.043	0.074	0.055	0.039
	Coverage	99	100	100	100	100	100	100	100	100	99
					Outcome 2: Paid assistive devices						
Comorbidity $\hat{\lambda}_{21}^{comp} = 0.086^{**}$	RMSE	0.016	0.028	0.013	0.020	0.019	0.016	0.020	0.021	0.109	0.011
	Width	0.051	0.054	0.057	0.052	0.048	0.05	0.049	0.051	0.256	0.049
	Coverage	87	44	96	81	76	92	75	79	78	98
BMI $\hat{\lambda}_{22}^{comp} = 0.003$	RMSE	0.003	0.002	0.003	0.002	0.002	0.002	0.003	0.002	0.005	0.003
	Width	0.012	0.014	0.012	0.012	0.012	0.012	0.013	0.011	0.021	0.013
	Coverage	100	100	99	100	100	100	97	97	99	99
					Correlation between the two outcomes						
$\hat{\rho} = 0.30$	RMSE	0.038	0.037	0.037	0.035	0.045	0.051	0.044	0.035	0.039	0.045
	Width	0.231	0.226	0.230	0.230	0.252	0.241	0.220	0.243	0.233	0.221
	Coverage	100	100	100	100	100	98	100	100	100	97

TABLE 7 The root mean squared error (RMSE), average interval width and empirical coverage of the 95% CI of coefficients estimates of the incomplete predictors and the correlation estimate for the joint analysis in Configuration 3.

Estimates	Missing pattern	FCS-Standard (LM)	FCS-Standard (PMM)	FCS-Standard (Poisson)	FCS-LMM-latent	FCS-GLMM (Gauss.)	FCS-GLMM (Poisson)	JM-General Location	JM-MLMM (common)	JM-MLMM (random)	JM-MGLMM
Intercept	Monotone	0.28	0.29	0.31	0.23	0.23	0.22	0.19	0.23	0.29	0.23
	Intermittent	0.32	0.30	0.31	0.26	0.27	0.26	0.17	0.25	0.35	0.23
Comorbidity Index	Monotone	0.26	0.27	0.51	0.23	0.23	0.37	0.19	0.23	0.36	0.38
	Intermittent	0.37	0.38	0.48	0.33	0.36	0.47	0.20	0.34	0.45	0.36
BMI	Monotone	0.28	0.27	0.30	0.24	0.24	0.23	0.19	0.23	0.33	0.23
	Intermittent	0.33	0.33	0.33	0.27	0.30	0.32	0.18	0.28	0.46	0.25
Paid assistive devices	Monotone	0.24	0.30	0.23	0.21	0.19	0.18	0.16	0.21	0.37	0.24
	Intermittent	0.49	0.51	0.47	0.41	0.40	0.39	0.26	0.37	0.48	0.31
Age	Monotone	0.27	0.25	0.28	0.21	0.24	0.24	0.17	0.22	0.23	0.22
	Intermittent	0.25	0.25	0.25	0.20	0.26	0.26	0.14	0.21	0.21	0.20
Gender	Monotone	0.27	0.26	0.28	0.21	0.24	0.24	0.18	0.22	0.22	0.21
	Intermittent	0.27	0.25	0.26	0.20	0.27	0.28	0.15	0.22	0.22	0.19
Self-rated health	Monotone	0.30	0.30	0.35	0.22	0.24	0.26	0.19	0.23	0.26	0.23
	Intermittent	0.27	0.27	0.30	0.23	0.27	0.29	0.16	0.24	0.28	0.21
Have med. bill paid off overtime	Monotone	0.23	0.23	0.23	0.17	0.22	0.22	0.16	0.20	0.20	0.20
	Intermittent	0.23	0.23	0.22	0.18	0.26	0.26	0.11	0.19	0.20	0.18
Took prescribed med.	Monotone	0.34	0.34	0.35	0.23	0.25	0.25	0.23	0.23	0.23	0.23
	Intermittent	0.36	0.35	0.35	0.26	0.28	0.29	0.20	0.29	0.31	0.23
Time	Monotone	0.32	0.32	0.36	0.26	0.30	0.30	0.21	0.29	0.28	0.25
	Intermittent	0.36	0.36	0.37	0.29	0.32	0.32	0.18	0.33	0.33	0.27
Subject-level variance	Monotone	0.36	0.35	0.37	0.33	0.22	0.23	0.24	0.35	0.35	0.28
	Intermittent	0.44	0.45	0.47	0.40	0.31	0.33	0.23	0.41	0.43	0.33

TABLE 8 The fraction of missing information estimated for the univariate GLMM analysis in Configuration 3.
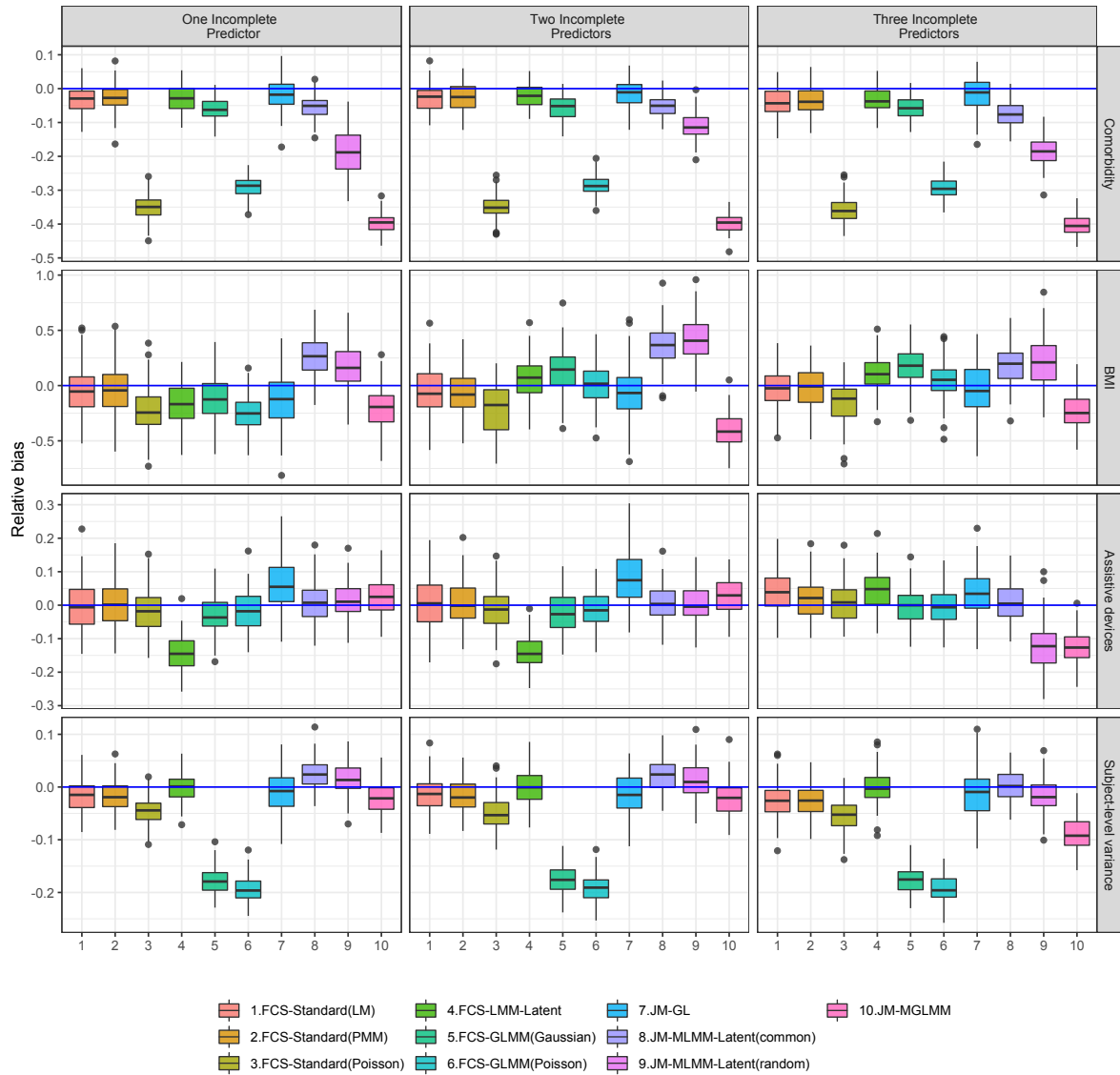
FIGURE 1 Relative bias of regression coefficients estimates associated with the incomplete covariates and the random effect estimate. Each column represents a simulation setting. From left to right, the column corresponds to Configuration 1 to Configuration 3.
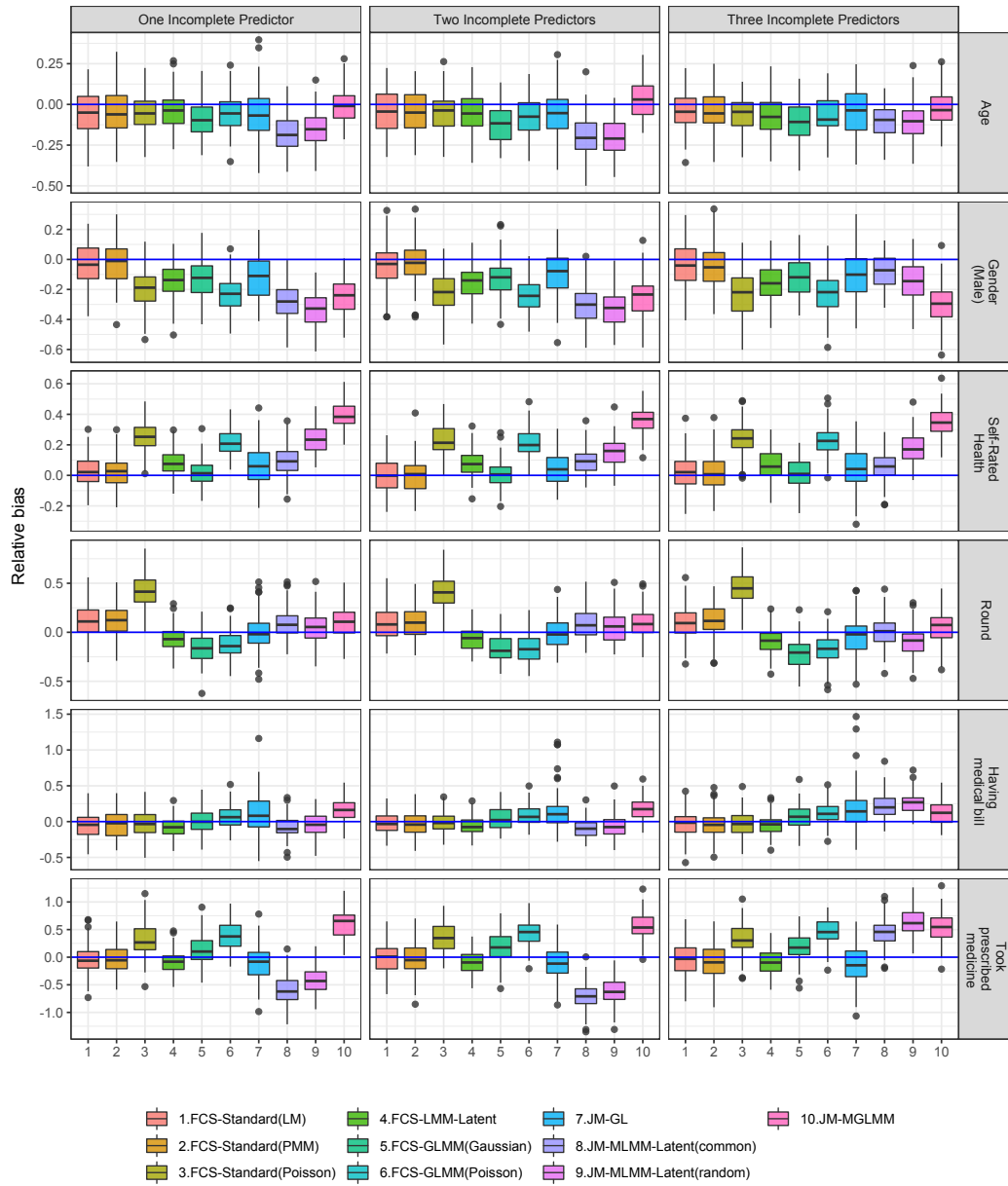
FIGURE 2 Relative bias of regression coefficients estimates associated with the fully observed predictors. Each column represents a simulation setting. From left to right, the column corresponds to Configuration 1 to Configuration 3.
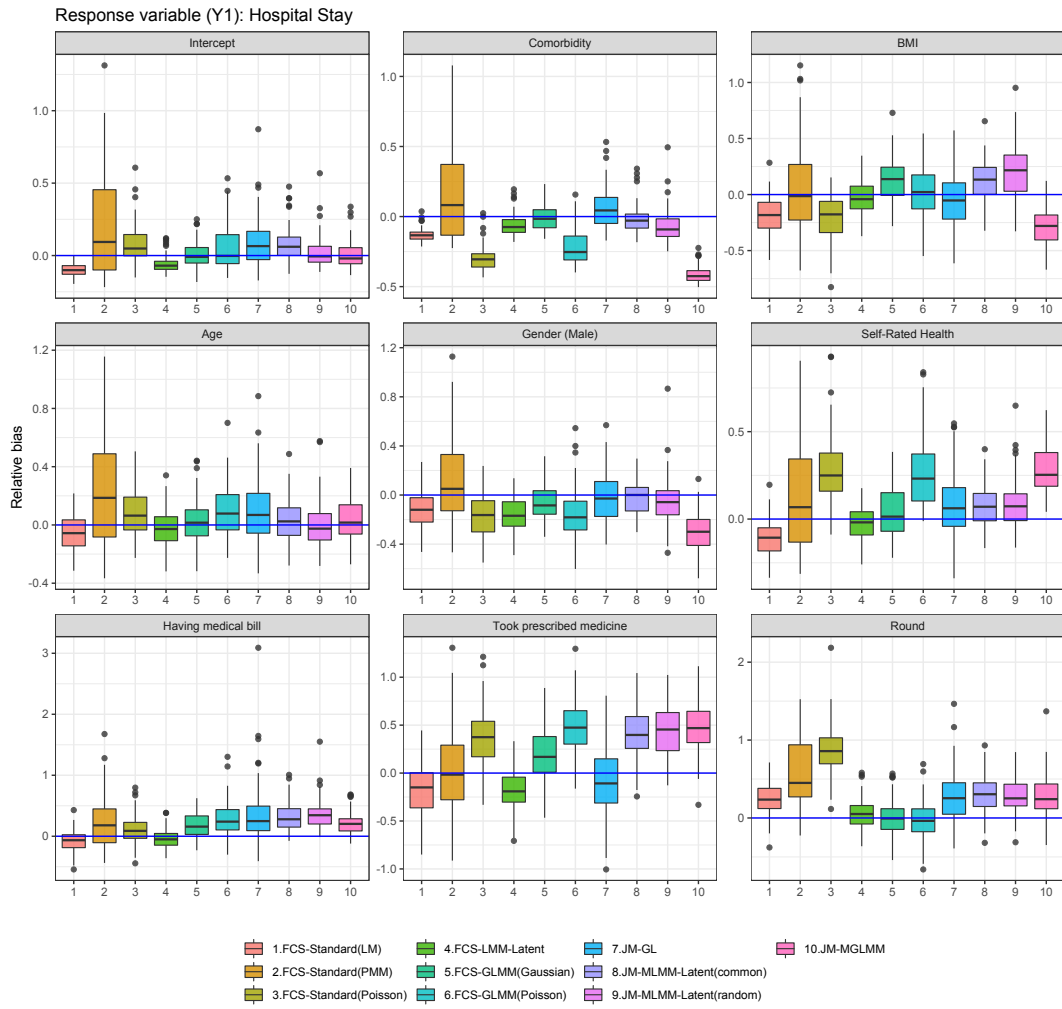
FIGURE 3 Relative bias of regression coefficients estimates for hospital stay.

FIGURE 4 Odds ratios estimates of sociodemographic factors, the number of hospital admissions, measures of physical and cognitive abilities on skilled nursing facilities admissions for different imputation methods.

APPENDIX

A SIMULATION RESULTS OF UNIVARIATE MODEL ANALYSIS FOR CONFIGURATIONS 3 WITH INTERMITTENT MISSING DATA PATTERN

B SIMULATION RESULTS OF UNIVARIATE MODEL ANALYSIS FOR CONFIGURATIONS 1 AND 2

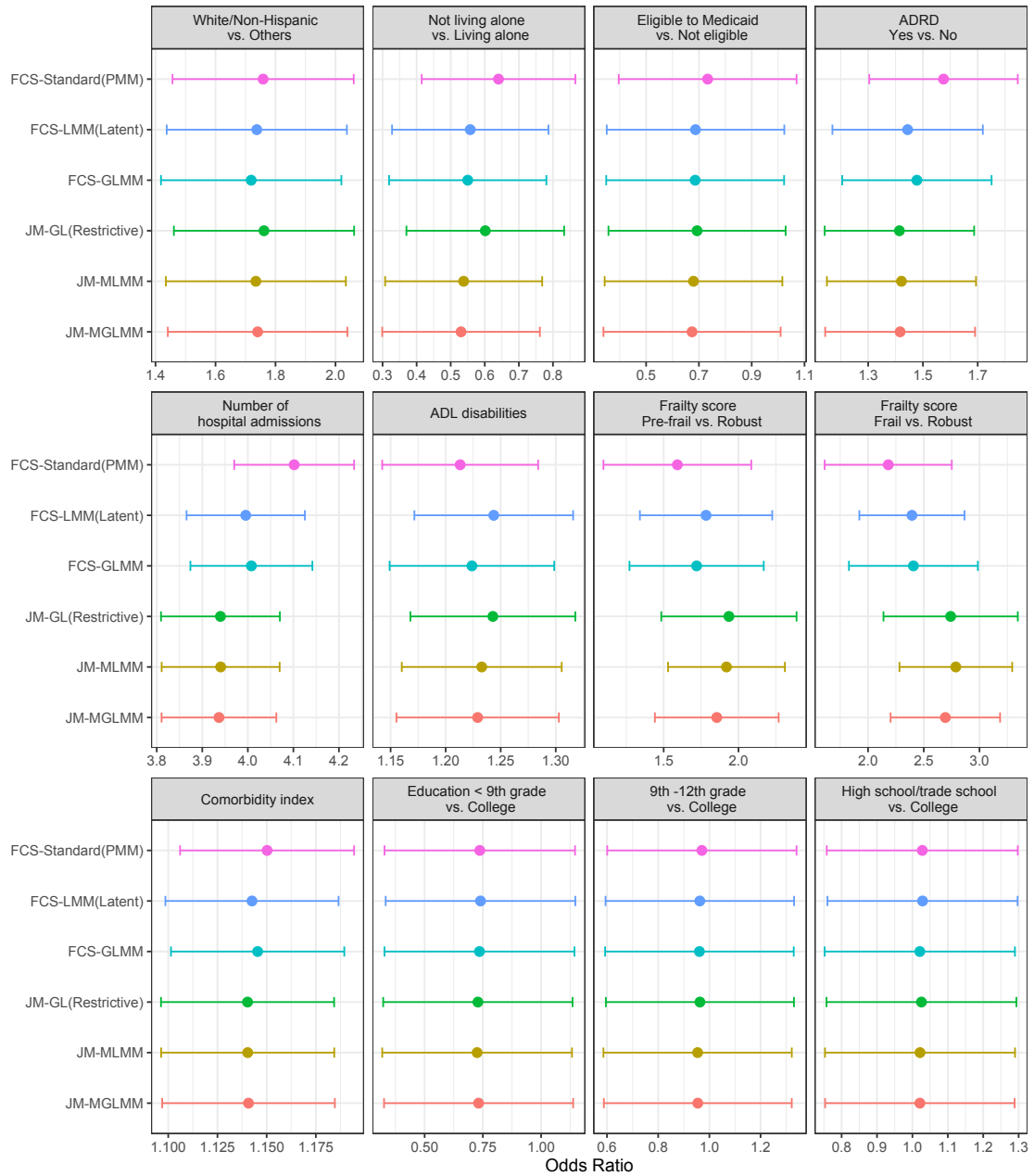| Estimates	Metrics	FCS-Standard (LM)	FCS-Standard (PMM)	FCS-Standard (Poisson)	FCS-LMM-latent	FCS-GLMM (Gauss.)	FCS-GLMM (Poisson)	JM-General Location	JM-MLMM (common)	JM-MLMM (random)	JM-MGLMM
Intercept $\hat{\gamma}_0^{comp} = -3.89^{**}$	MRB	-0.018	-0.031	0.080	-0.045	0.039	0.086	0.153	0.101	0.170	0.224
	RMSE	0.170	0.176	0.166	0.201	0.236	0.219	0.252	0.480	0.535	0.255
	Width	1.170	1.154	1.149	1.123	1.027	1.009	1.075	1.116	1.170	1.026
	Coverage(%)	100	100	100	99	100	99	97	76	67	99
Comorbidity Index $\hat{\gamma}_1^{comp} = 0.38^{**}$	MRB	0.046	0.011	0.027	-0.099	0.002	-0.027	0.050	-0.001	-0.197	-0.332
	RMSE	0.026	0.029	0.152	0.046	0.060	0.141	0.027	0.041	0.116	0.239
	Width	0.127	0.127	0.115	0.122	0.112	0.106	0.113	0.124	0.120	0.089
	Coverage(%)	98	96	0	85	44	0	95	90	3	0
BMI $\hat{\gamma}_2^{comp} = -0.02^*$	MRB	-0.036	-0.050	-0.401	-0.112	-0.153	-0.371	-0.017	-0.099	-0.301	-0.634
	RMSE	0.004	0.004	0.005	0.004	0.004	0.006	0.005	0.010	0.010	0.011
	Width	0.030	0.030	0.029	0.029	0.026	0.026	0.027	0.029	0.033	0.024
	Coverage(%)	100	100	99	100	100	99	99	94	98	62
Paid assistive devices $\hat{\gamma}_3^{comp} = 0.36^{**}$	MRB	0.031	0.089	0.185	-0.089	-0.247	-0.154	-0.140	0.085	-0.126	0.120
	RMSE	0.041	0.036	0.035	0.043	0.025	0.025	0.048	0.026	0.077	0.122
	Width	0.183	0.185	0.181	0.172	0.159	0.158	0.154	0.167	0.167	0.146
	Coverage(%)	96	99	100	94	100	100	88	99	63	1
Age $\hat{\gamma}_4^{comp} = 0.11^{**}$	MRB	-0.013	-0.018	-0.001	-0.043	-0.053	-0.047	-0.007	-0.120	-0.132	0.057
	RMSE	0.015	0.015	0.014	0.025	0.017	0.014	0.017	0.029	0.025	0.016
	Width	0.111	0.111	0.110	0.109	0.098	0.098	0.106	0.110	0.108	0.102
	Coverage(%)	100	100	100	100	100	100	100	99	99	100
Gender $\hat{\gamma}_5^{comp} = 0.28^{**}$	MRB	-0.051	-0.052	-0.241	-0.254	-0.212	-0.298	-0.088	-0.268	-0.350	-0.458
	RMSE	0.040	0.041	0.077	0.076	0.067	0.088	0.052	0.081	0.102	0.132
	Width	0.311	0.307	0.304	0.299	0.272	0.272	0.293	0.305	0.301	0.280
	Coverage(%)	100	100	98	100	99	94	100	99	95	62
Self-rated health $\hat{\gamma}_6^{comp} = 0.20^{**}$	MRB	-0.002	0.011	0.354	0.000	0.000	0.217	0.014	0.068	0.287	0.668
	RMSE	0.022	0.022	0.073	0.018	0.019	0.047	0.026	0.023	0.061	0.135
	Width	0.160	0.160	0.159	0.157	0.141	0.140	0.151	0.159	0.161	0.141
	Coverage(%)	100	100	66	100	100	94	99	100	87	0
Have med. bill paid off overtime $\hat{\gamma}_7^{comp} = 0.34^*$	MRB	-0.035	-0.046	0.010	-0.204	-0.108	-0.072	-0.033	-0.233	-0.195	0.092
	RMSE	0.059	0.059	0.062	0.048	0.051	0.055	0.093	0.059	0.076	0.091
	Width	0.481	0.481	0.473	0.469	0.424	0.419	0.453	0.478	0.473	0.444
	Coverage(%)	100	100	100	100	100	100	97	100	100	100
Took prescribed med. $\hat{\gamma}_8^{comp} = 0.30$	MRB	0.031	0.039	0.429	0.081	0.261	0.495	-0.129	-0.247	0.081	1.055
	RMSE	0.104	0.101	0.165	0.065	0.111	0.168	0.215	0.105	0.081	0.331
	Width	0.722	0.722	0.710	0.677	0.632	0.630	0.658	0.682	0.683	0.637
	Coverage(%)	100	99	98	100	100	99	97	100	100	47
Time $\hat{\gamma}_9^{comp} = 0.10^{**}$	MRB	0.009	0.016	-0.191	-0.091	-0.143	-0.260	-0.083	0.510	0.473	-0.594
	RMSE	0.017	0.018	0.025	0.016	0.029	0.021	0.024	0.017	0.020	0.019
	Width	0.110	0.111	0.110	0.105	0.105	0.104	0.098	0.109	0.108	0.102
	Coverage(%)	100	100	99	100	95	99	97	100	99	100
Random effect $\hat{\sigma}_{b_{comp}}^2 = 1.13^{**}$	MRB	-0.027	-0.027	-0.041	-0.018	-0.297	-0.305	-0.010	0.007	-0.017	-0.099
	RMSE	0.053	0.055	0.063	0.042	0.338	0.346	0.067	0.038	0.042	0.120
	Width	0.256	0.257	0.257	0.247	0.232	0.232	0.219	0.250	0.253	0.225
	Coverage(%)	98	97	96	100	0	0	90	100	100	53

TABLE A1 The mean relative bias (MRB), the root mean squared error (RMSE), average interval width and empirical coverage of the 95% CI of coefficients estimates and subject-level variance estimate for the univariate analysis in Configuration 3 with intermittent missing data pattern. The incomplete predictors are marked in bold.

Estimates	Metrics	FCS-Standard (LM)	FCS-Standard (PMM)	FCS-Standard (Poisson)	FCS-LMM-latent	FCS-GLMM (Gauss.)	FCS-GLMM (Poisson)	JM-General Location	JM-MLMM (common)	JM-MLMM (random)	JM-MGLMM
Intercept $\hat{\gamma}_0^{comp} = -3.89^{**}$	RMSE	0.133	0.134	0.144	0.115	0.164	0.145	0.159	0.384	0.349	0.153
	Width	1.156	1.143	1.125	1.119	1.047	1.041	1.078	1.116	1.116	1.121
	Coverage	100	100	100	100	100	100	100	91	96	100
Comorbidity Index $\hat{\gamma}_1^{comp} = 0.38^{**}$	RMSE	0.020	0.018	0.132	0.018	0.027	0.110	0.018	0.024	0.075	0.150
	Width	0.117	0.118	0.117	0.115	0.108	0.106	0.114	0.115	0.185	0.101
	Coverage	100	99	0	100	98	0	99	100	75	0
BMI $\hat{\gamma}_2^{comp} = -0.02^{*}$	RMSE	0.004	0.004	0.006	0.004	0.004	0.005	0.005	0.006	0.005	0.005
	Width	0.029	0.028	0.028	0.028	0.026	0.026	0.027	0.028	0.028	0.028
	Coverage	100	100	99	100	100	100	99	100	100	100
Paid assistive devices $\hat{\gamma}_3^{comp} = 0.36^{**}$	RMSE	0.026	0.024	0.024	0.055	0.023	0.022	0.035	0.022	0.021	0.022
	Width	0.151	0.153	0.149	0.147	0.142	0.141	0.141	0.150	0.148	0.148
	Coverage	99	100	100	84	100	100	96	100	100	100
Age $\hat{\gamma}_4^{comp} = 0.11^{**}$	RMSE	0.015	0.016	0.015	0.014	0.016	0.014	0.018	0.024	0.022	0.013
	Width	0.113	0.113	0.111	0.110	0.103	0.099	0.106	0.113	0.111	0.109
	Coverage	100	100	100	100	100	100	100	100	100	100
Gender $\hat{\gamma}_5^{comp} = 0.28^{**}$	RMSE	0.039	0.040	0.065	0.052	0.050	0.072	0.053	0.084	0.099	0.077
	Width	0.31	0.309	0.305	0.305	0.287	0.279	0.295	0.309	0.307	0.302
	Coverage	100	100	100	100	100	99	100	99	98	99
Self-rated health $\hat{\gamma}_6^{comp} = 0.20^{**}$	RMSE	0.021	0.021	0.053	0.023	0.016	0.046	0.028	0.027	0.051	0.081
	Width	0.161	0.163	0.159	0.159	0.147	0.144	0.155	0.160	0.178	0.157
	Coverage	100	100	96	100	100	97	99	100	94	49
Have med. bill paid off overtime $\hat{\gamma}_7^{comp} = 0.34^{*}$	RMSE	0.065	0.064	0.065	0.054	0.058	0.060	0.099	0.058	0.060	0.078
	Width	0.484	0.479	0.478	0.475	0.443	0.436	0.461	0.488	0.481	0.471
	Coverage	100	100	100	100	100	100	98	100	100	100
Took prescribed med. $\hat{\gamma}_8^{comp} = 0.30$	RMSE	0.078	0.084	0.135	0.066	0.089	0.144	0.099	0.197	0.148	0.197
	Width	0.701	0.724	0.689	0.674	0.648	0.638	0.669	0.686	0.697	0.674
	Coverage	100	100	99	100	100	100	100	97	100	99
Time $\hat{\gamma}_9^{comp} = 0.10^{**}$	RMSE	0.02	0.021	0.046	0.015	0.022	0.020	0.018	0.018	0.017	0.019
	Width	0.104	0.107	0.112	0.103	0.101	0.103	0.099	0.104	0.108	0.104
	Coverage	100	99	75	100	99	100	98	99	99	98
Random effect $\hat{\sigma}_{b_{comp}}^2 = 1.13^{**}$	RMSE	0.042	0.040	0.060	0.029	0.204	0.222	0.044	0.041	0.037	0.041
	Width	0.239	0.237	0.240	0.237	0.213	0.218	0.218	0.237	0.236	0.237
	Coverage	100	100	99	100	0	0	99	99	100	100

TABLE B2 The root mean squared error (RMSE), average interval width and empirical coverage of the 95% CI of coefficients estimates and subject-level variance estimate for the univariate analysis in Configuration 1. The incomplete predictors are marked in bold.

Estimates	Metrics	FCS-Standard (LM)	FCS-Standard (PMM)	FCS-Standard (Poisson)	FCS-LMM-latent)	FCS-GLMM (Gauss.)	FCS-GLMM (Poisson)	JM-General Location	JM-MLMM (common)	JM-MLMM (random)	JM-MGLMM
Intercept $\hat{\gamma}_0^{comp} = -3.89^{**}$	RMSE	0.138	0.140	0.139	0.154	0.239	0.228	0.175	0.483	0.520	0.222
	Width	1.167	1.131	1.129	1.125	1.047	1.027	1.096	1.111	1.134	1.117
	Coverage	100	100	100	100	98	100	99	71	67	100
Comorbidity Index $\hat{\gamma}_1^{comp} = 0.38^{**}$	RMSE	0.018	0.019	0.133	0.015	0.025	0.108	0.017	0.023	0.044	0.15
	Width	0.119	0.118	0.116	0.117	0.108	0.103	0.113	0.115	0.130	0.102
	Coverage	100	100	0	100	99	0	100	100	89	0
BMI $\hat{\gamma}_2^{comp} = -0.02^{*}$	RMSE	0.004	0.004	0.005	0.003	0.004	0.003	0.004	0.007	0.008	0.008
	Width	0.029	0.028	0.028	0.028	0.026	0.026	0.027	0.029	0.030	0.027
	Coverage	100	100	100	100	99	100	100	99	97	98
Paid assistive devices $\hat{\gamma}_3^{comp} = 0.36^{**}$	RMSE	0.024	0.022	0.023	0.054	0.022	0.020	0.041	0.020	0.019	0.022
	Width	0.150	0.150	0.149	0.147	0.145	0.143	0.143	0.149	0.152	0.15
	Coverage	100	100	100	91	100	100	93	100	100	100
Age $\hat{\gamma}_4^{comp} = 0.11^{**}$	RMSE	0.016	0.015	0.015	0.015	0.019	0.016	0.017	0.027	0.026	0.013
	Width	0.113	0.112	0.112	0.112	0.103	0.101	0.107	0.111	0.112	0.109
	Coverage	100	100	100	100	100	100	100	99	100	100
Gender $\hat{\gamma}_5^{comp} = 0.28^{**}$	RMSE	0.041	0.039	0.071	0.053	0.049	0.074	0.047	0.090	0.096	0.078
	Width	0.313	0.315	0.306	0.304	0.285	0.279	0.298	0.314	0.304	0.303
	Coverage	100	100	99	100	100	98	99	98	96	98
Self-rated health $\hat{\gamma}_6^{comp} = 0.20^{**}$	RMSE	0.021	0.023	0.051	0.023	0.017	0.045	0.025	0.025	0.036	0.074
	Width	0.162	0.163	0.162	0.159	0.15	0.146	0.154	0.158	0.161	0.156
	Coverage	100	100	93	100	100	97	100	100	99	55
Have med. bill paid off overtime $\hat{\gamma}_7^{comp} = 0.34^{*}$	RMSE	0.051	0.058	0.049	0.050	0.053	0.055	0.096	0.055	0.058	0.077
	Width	0.485	0.493	0.476	0.473	0.435	0.440	0.459	0.480	0.485	0.466
	Coverage	100	100	100	100	100	100	96	100	100	100
Took prescribed med. $\hat{\gamma}_8^{comp} = 0.30$	RMSE	0.084	0.087	0.138	0.067	0.096	0.153	0.130	0.229	0.205	0.189
	Width	0.731	0.708	0.703	0.667	0.654	0.639	0.657	0.698	0.687	0.676
	Coverage	100	100	100	100	100	100	99	95	95	99
Time $\hat{\gamma}_9^{comp} = 0.10^{**}$	RMSE	0.019	0.018	0.044	0.014	0.022	0.022	0.017	0.017	0.017	0.017
	Width	0.106	0.105	0.110	0.104	0.104	0.103	0.099	0.104	0.104	0.105
	Coverage	99	100	80	100	100	100	100	100	100	100
Random effect $\hat{\sigma}_{b_{comp}}^2 = 1.13^{**}$	RMSE	0.041	0.040	0.065	0.036	0.200	0.218	0.047	0.043	0.041	0.046
	Width	0.237	0.242	0.235	0.240	0.214	0.215	0.219	0.236	0.240	0.236
	Coverage	100	100	96	100	0	0	97	100	99	100

TABLE B3 The root mean squared error (RMSE), average interval width and empirical coverage of the 95% CI of coefficients estimates and subject-level variance estimate for the univariate analysis in Configuration 2. The incomplete predictors are marked in bold.