Research Article

Multiple Imputation of Missing Covariates in NONMEM and Evaluation of the Method's Sensitivity to η-Shrinkage

Åsa M. Johansson[1,2] and Mats O. Karlsson[1]

Abstract. Multiple imputation (MI) is an approach widely used in statistical analysis of incomplete data. However, its application to missing data problems in nonlinear mixed-effects modelling is limited. The objective was to implement a four-step MI method for handling missing covariate data in NONMEM and to evaluate the method's sensitivity to η-shrinkage. Four steps were needed; (1) estimation of empirical Bayes estimates (EBEs) using a base model without the partly missing covariate, (2) a regression model for the covariate values given the EBEs from subjects with covariate information, (3) imputation of covariates using the regression model and (4) estimation of the population model. Steps (3) and (4) were repeated several times. The procedure was automated in PsN and is now available as the mimp functionality (http://psn.sourceforge.net/). The method's sensitivity to shrinkage in EBEs was evaluated in a simulation study where the covariate was missing according to a missing at random type of missing data mechanism. The η-shrinkage was increased in steps from 4.5 to 54%. Two hundred datasets were simulated and analysed for each scenario. When shrinkage was low the MI method gave unbiased and precise estimates of all population parameters. With increased shrinkage the estimates became less precise but remained unbiased.

KEY WORDS: covariates; missing data; multiple imputation; NONMEM.

INTRODUCTION

Missing data is a frequently encountered problem which needs to be considered in all statistical analyses (1). The statistical power is reduced with decreasing amount of data and it is therefore important to retain as much data as possible. Various imputation methods have been suggested to fill in missing values in order to receive complete data sets. Since variables in the data set might contain partly the same information (2), relationships can be established between the observed and the missing data and the imputations can be based on these relationships (3,4). When the imputations are done only once (single imputation) the completed data set will be analysed as if the imputed values are the true ones, without taking the uncertainty in the established relationships into account. Rubin suggests repeating the imputations many times (multiple imputation), by sampling from the predictive distribution of the missing data given the observed data, and then analyse each completed data set and combine the statistics considering the variability (5,6). There are several MI methods available for handling missing data in linear mixed effects models (7–9) and the Journal of Statistical Software give a survey of software implementations of MI methods in their December 2011 issue (10–15). However, so far there is only one MI method available which can handle missing data in general nonlinear mixed effects modelling (16).

Longitudinal clinical data are usually analysed with nonlinear mixed effects modelling to characterize the pharmacokinetic (PK) and/or pharmacodynamic (PD) properties of the investigated treatment, where the fixed effects are the typical values in the population and the random effects describe the between subject parameter variability. The predictability of the models is improved by inclusion of covariates, such as demographics and physiological information, which explain some of the observed between subject variability. Information about missing covariates can be available in the observed covariates and in the dependent variable (response variable) and proper imputations should therefore be based on relationships established between these variables (17–20). Since the model for the dependent variable is nonlinear with multiple hierarchies of random effects it cannot be used directly in the imputation model. Wu and Wu suggest a MI method where the imputations are based on observed covariates and individual parameter estimates (16). The individual parameter estimates contain information about the dependent variable (21,22) and hence also about the missing covariates. With decreasing information on the individual level in the data set to be analysed, the individual parameter estimates will shrink towards the estimates of the fixed effects (η-shrinkage) (23). Shrunken individual estimates will affect the imputed values and it is therefore important to investigate how this will influence the final model. Wu and Wu implemented their MI method in a hierarchical Bayesian setting while in this study the method was implemented in a maximum likelihood model.

NONMEM (24) is the most widely used software in nonlinear mixed effects modelling of PK/PD data. The objective of this study was to implement the MI method presented by Wu

[1] Department of Pharmaceutical Biosciences, Uppsala University, P.O. Box 591, 751 24 Uppsala, Sweden.
[2] To whom correspondence should be addressed. (e-mail: asa.johansson@farmbio.uu.se)

and Wu in NONMEM and to evaluate the method's sensitivity to η-shrinkage.

METHODS

The Multiple Imputation Method

General Model

Let N be the number of subjects in the study, independently sampled from the study population, and let y_i be a vector of the observed dependent variable (*e.g.* measured concentrations) of individual i ($i=1,...,N$). Then the general nonlinear mixed effects model can be defined as

$$y_i = f(t_i, g(x_i, z_i, \theta_i)) + h(t_i, g(x_i, z_i, \theta_i), \varepsilon_i) \qquad (1)$$

where $f()$ is the vector function of the model, $h()$ is the vector function describing the residual error model, t_i is a vector of the independent variable (*e.g.* observation times) of size n_i (number of independent variables of individual i) and $g()$ is a vector function defining the relationship between the vector of discrete design variables x_i (*e.g.* dose), the vector of covariates z_i (*e.g.* weight), and the vector of individual parameters θ_i, of individual i. The individual parameters of the ith individual deviate from the population fixed effects θ with the random effects η_i ($\boldsymbol{\eta} \sim N(0,\Omega)$), where θ and η_i are vectors of size s and Ω is the $s \times s$ covariance matrix describing the correlations between the individual parameters. The diagonal elements of Ω are also referred to as the between subject variabilities. In the residual error model ε_i ($\boldsymbol{\varepsilon} \sim N(0,\Sigma)$) describes the deviation of the model predictions from the observed values of the dependent variable, where ε_i is a vector of size n_i and Σ is the covariance matrix describing the correlation between the $\sum_{i=1}^{N} n_i$ residual error parameters.

Missing Covariates

The number of observed covariates can vary from individual to individual. Suppose that p is the total number of covariates intended to be observed according to the study design and let $p=q+r$, where q is the number of covariates which are completely observed for all individuals and r is the number of covariates which are incompletely observed. $\mathbf{Z}$ is the $N \times p$ matrix of all covariates intended to be observed and the ith row of the matrix is the vector z_i. The $\mathbf{Z}$ matrix can be divided into $\mathbf{U}$ and $\mathbf{V}$, where $\mathbf{U}$ is the $N \times q$ matrix containing the completely observed covariates and $\mathbf{V}$ is the $N \times r$ matrix containing the incompletely observed covariates. The ith row of $\mathbf{U}$ is the vector $u_i=(u_{i1},...,u_{iq})$ and the ith row of $\mathbf{V}$ is the vector $v_i=(v_{i1},...,v_{ir})$. $\mathbf{B}$ is a binary $N \times r$ matrix that indicates which data in $\mathbf{V}$ are missing (1=missing, 0=observed). The nonlinear mixed effects model in Eq. 1 can be directly estimated if only the subjects in N for whom all data are available are kept in the analysis or if z_i is exchanged to u_i. Both scenarios imply that potentially informative data are excluded from the analysis.

Imputation Model

Wu and Wu suggest the following four steps MI method where the information available in the observed covariates and the dependent variable is used to impute the values of the missing covariates (16);

Step 1 Estimate the parameters in (1) without inclusion of any covariates, alternatively only include covariates in $\mathbf{U}$. The model can then be written as

$$y_i = f\left(t_i, g\left(x_i, \theta_i^{(0)}\right)\right) + h\left(t_i, g\left(x_i, \theta_i^{(0)}\right), \varepsilon_i\right) \qquad (2)$$

where $\theta_i^{(0)}$ is the vector (size s) of individual parameters for individual i.

Step 2 Create an imputation model for the missing covariate values given the observed covariates and the individual parameter estimates ($\widehat{\theta}_i^{(0)}$) obtained from (2).

$$\mathbf{V} = k\left(\left(l\,\mathbf{U}\,\widehat{\theta}^{(0)}\right)\mathbf{D}, e\right) = k(\mathbf{WD}, e) \qquad (3)$$

where k is a matrix function describing the residual error model (a regression model), l represents a vector of size N which only contain ones ($l=(1,...,1)$), $\mathbf{W} = \left(l\,\mathbf{U}\,\widehat{\theta}^{(0)}\right)$ is an $N \times (1+q+s)$ matrix with the ith row $w_i = \left(1, u_{i1},...,u_{iq}, \widehat{\theta}_{i1}^{(0)},...,\widehat{\theta}_{is}^{(0)}\right)$, $\widehat{\theta}^{(0)}$ is an $N \times s$ matrix with individual parameter estimates of all parameters for all individuals, $\mathbf{D}$ is a $(1+q+s) \times r$ matrix of parameters describing the relationship between the observed covariates in $\mathbf{U}$ and the individual parameter estimates in $\widehat{\theta}^{(0)}$ with the partly missing covariates in $\mathbf{V}$, and e is an $N \times r$ matrix of residual errors $e \sim N(0,\Xi)$ where Ξ is a covariance matrix describing the correlation between the residual error parameters in e. All values in $\mathbf{W}$ and part of the values in $\mathbf{V}$ are known which enables estimation of the parameters in $\mathbf{D}$ and Ξ.

Step 3 The missing values in $\mathbf{V}$ are imputed by drawing m independent samples of each missing value using the model in (3) and the parameters estimated in step 2. Note that the samples will be drawn from the distributions defined by Ξ.

Step 4 Fit the model in (1) to the data in the m imputed data sets and combine the estimates according to the equations in (4) (6,19).

$$\tilde{\beta} = \frac{1}{m} \sum_{\gamma=1}^{m} \widehat{\beta}^{(\gamma)} \qquad (4)$$

$$\tilde{b} = \frac{1}{m} \sum_{\gamma=1}^{m} \widehat{b}^{(\gamma)} + \frac{m+1}{m(m-1)} \sum_{\gamma=1}^{m} \left(\widehat{\beta}^{(\gamma)} - \tilde{\beta}\right)^2$$

where β is a parameter in θ, Ω or Σ, $\tilde{\beta}$ is the calculated estimate of β, $\tilde{b}$ is the calculated variance associated with $\tilde{\beta}$, $\widehat{\beta}^{(\gamma)}$ is the estimate of β from imputed data set γ ($\gamma=1,...,m$) and $\widehat{b}^{(\gamma)}$ is the estimate of b from the imputed data set γ.

Shrinkage

The individual estimates generated in step 1 can suffer from η-shrinkage (23). The amount of shrinkage in the individual estimates for a particular individual is dependent on the magnitude of the variability of the residual errors (Σ), the

number of observations of the dependent variable for that individual and the informativeness of these observations (25). An increasing shrinkage will decrease the apparent impact of the partly missing covariate on the dependent variable. The weakened relationship will then be incorporated in the **D** matrix where it is used to impute the missing values in **V**. The shrinkage (sh) is evaluated by comparing the distribution of the individual estimates ($\widehat{\eta}$) with the corresponding estimated distribution in $\widehat{\Omega}$ for each parameter j ($j=1,\ldots,s$) using the empirical equation in (5) (25,26).

$$\mathrm{sh}_{\eta_j} = 1 - \frac{\mathrm{sd}\left(\widehat{\eta}_{ij}\right)}{\widehat{\omega}_j} \qquad (5)$$

where $\mathrm{sd}\left(\widehat{\eta}_{ij}\right)$ is the standard deviation of the N individual estimates of parameter j and $\widehat{\omega}_j$ is the square root of the jth diagonal element in $\widehat{\Omega}$.

Multiple Imputation in NONMEM

Three NONMEM models were developed to adapt the presented MI method to NONMEM syntax; a base model for estimation of $\widehat{\theta}^{(0)}$, a regression model for estimation of **D** and Ξ, and an imputation model for imputing the missing values in **V** and for fitting the model in (1) to the imputed data sets. Information was transmitted from one NONMEM model to the subsequent model(s) by printing the generated information in a table file which was then used as input to the next model.

NONMEM data sets are constructed as matrices and consist of data records (rows) and data items (columns) (24). There is one data record for each event (*e.g.* dosing event or sampling of dependent variable) occurring for each individual. All events associated with an individual are assembled in the consecutive data records of the matrix and the events are ordered according to the time they occurred. All data records contain the same number of data items. Data items which are only measured once during a study (*e.g.* non time varying covariates) are filled in so that all data records for an individual contain the same value of that data item. The following paragraph explains how the original NONMEM data set had to be modified to enable MI in NONMEM, how the dataset evolved when generated information was incorporated and how the information in the data set was used during estimation of parameters in the different NONMEM models.

The information in the binary matrix **B** was added to the original data set as new data items. For all data records where the partly missing covariate(s) was observed the corresponding new data item was assigned zeros (0) and for all data records where the covariate(s) was missing the corresponding new data item was assigned ones (1). The new data set (data1) was used as input to the base model. The individual parameter estimates in $\widehat{\theta}^{(0)}$ were obtained after fitting the base model (Eq. 2) to data1. The estimates in $\widehat{\theta}^{(0)}$ were added to data1 as new data item(s) (same individual parameter estimate(s) in all data records of an individual) and the extended data set (data2) was generated and outputted from the base model. The file data2 was used as input to the regression model. In the regression model the data item of the partly missing covariate

functioned as the dependent variable and the corresponding binary data item (which indicated when the partly missing covariate was observed and when it was not) served as a missing dependent variable data item. Only data records where the dependent variable was observed (*i.e.* where the missing dependent variable data item was '0') was used during the estimation of the regression model parameters in **D** and Ξ. Note that this step was executed for one covariate at a time which yielded a **D** matrix of size $(1+q+s)\times1$ and an **e** matrix of size $N\times1$ (*i.e.* **D** and **e** were column vectors). The estimated parameters in **D** ($\widehat{\mathbf{D}}$) and Ξ ($\widehat{\Xi}$) were added to data2 as new data items (same parameter estimates in all data records of all individuals) and the extended data set (data3) was generated and outputted from the regression model. The file data3 was used as input to the imputation model. For each individual with a missing covariate the missing values were imputed by using the estimates in $\widehat{\mathbf{D}}$ and $\widehat{\Xi}$ together with the value(s) of the individuals' observed covariate(s) and estimated individual parameter ($\widehat{\theta}^{(0)}$) and by drawing (*i.e.* simulating) a random value from $\widehat{\Xi}$. The model in (1) was fitted to the imputed data set directly in the imputation model. The imputation step could be repeated by drawing new random values from the residual error distribution.

The MI procedure was automated in PsN (version 3.5.3) (27,28) and is available as the mimp functionality. The user can choose number of imputations and there is a possibility to add extra models which will be estimated between the base model and the regression model. The mimp functionality can also be used in a stochastic simulation and estimation (SSE) type of setting and in that case a simulation model has to be provided.

Evaluation of the Multiple Imputation Method

An SSE analysis was applied to evaluate the implemented MI method's performance with respect to bias and precision in population parameter estimates and to evaluate the method's sensitivity to η-shrinkage. Simulations and model analyses were performed in NONMEM 7.1.2 facilitated with PsN 3.5.3. All models were fitted to data using the first-order conditional estimation algorithm, except the regression model in (3) which was fitted using the first-order algorithm. Statistical analyses of the data were completed in R 2.14.1 (http://www.r-project.org).

Population Model

A population pharmacokinetic (PK) model with constant infusion at steady state was used for simulations and estimations. A covariate effect was implemented on clearance (CL) and two levels of the effect were investigated. The relative difference in the fixed effect of CL between males and females were 17 or 50%, where females were assigned to have a lower CL than males. The individual CL values were log-normally distributed around the fixed effects with a random effect of 30% and the residual error model was proportional with a random effect of 20, 50 or 70%.

Simulation of Covariates

Each data set consisted of data for 200 individuals. The individuals were randomly assigned to be males (60%) or females (40%) and their weights were simulated from truncated log-normal distributions ($\ln N$(85.1, 0.0329) for males and $\ln N$

(73.0, 0.0442) for females). The fixed and random effects of the sex-specific weight distributions were estimated, prior to the simulations, using an external dataset with 1,022 males and 423 females (29). The number of observations (*i.e.* concentration measurements) simulated was either two for all individuals or two for one third of the individuals and one for the remaining two thirds (equally distributed between males and females). In each data set 50% of the individuals were assigned to lack information about the covariate sex according to a missing at random type of missing data mechanism (1). The mechanism gave a higher probability of missing sex with increasing weight, *e.g.* the probability of missing information about sex was 27% for an individual weighing 40 kg and 83% for an individual weighing 145 kg. The proportion of males among the individuals with observed sex was approximately 56%.

Multiple Imputation of the Missing Covariate

Individual estimates of CL were obtained from the base model and the relationship between weight, the individual estimate of CL and sex was estimated as a logistic regression in the regression model. In the imputation model the missing sexes were imputed by drawing (simulating) random values from the uniform distribution [0, 1] and comparing these values with each individual's probability of being male given by the logistic regression curve. When the drawn value was lower than or equal to the estimated probability the missing sex was imputed as 'male' otherwise 'female'. After imputation the population model was fitted to the imputed data set. The imputation followed by estimation was repeated six times for each simulated dataset and the estimates reported were calculated according to the equations in (4).

Simulation Study

A total of eight scenarios were investigated, where the settings were altered to change the amount of information available on the individual level in the simulated data sets (Table I, columns 2–4). An SSE analysis was conducted where 200 data sets were simulated for each scenario. The data sets were analysed using the MI method and, to enable a comparison, they were also analysed using all data. The difference in objective function values (ΔOFV) between the base model and the model including the covariate effect was considered approximately χ^2-distributed and a decrease of at least 3.84 in OFV when adding the covariate effect (one extra parameter) was regarded as significant ($p < 0.05$). η-shrinkage is theoretically independent on number of individuals in the data set. However, a greater number of individuals will reduce the uncertainty in the estimates of θ and Ω and hence improve the precision of the estimated η-shrinkage when calculated using the empirical equation (Eq. 5) (25). The level of η-shrinkage was calculated using Eq. 5 after simulating data for 10,000 individuals (6,000 males and 4,000 females) for each scenario and fitting the base model to the data.

Comparison of Bias and Precision

The efficiency of the MI method was evaluated by comparing the bias and precision in the estimated fixed and random effects with the bias and precision received when all data

was used in the estimations. The bias was defined as the deviation of the mean of the estimates from the true value and the relative bias (RBias) was hence calculated as;

$$\text{RBias}\left[\widetilde{\beta}\right] = \frac{\beta - \beta_T}{\beta_T} \tag{6}$$

where β is the parameter (fixed or random effect), $\widetilde{\beta}$ is the calculated estimate of β (see Eq. 4), $\overline{\widetilde{\beta}}$ is the mean of the estimates of β and β_T is the true value of the parameter, *i.e.* the value used in the simulation. A RBias <5% for the fixed effect parameters and <10% for the random effect parameters was considered as no bias.

The precision was defined as the spread of the estimates around the mean estimate ($\overline{\beta}$) of β and was evaluated by calculating the relative standard deviation (RSD);

$$\text{sd}\left[\widetilde{\beta}\right] = \sqrt{\frac{1}{C-1} \sum_{\varphi=1}^{C} \left(\widetilde{\beta}_\varphi - \overline{\beta}\right)^2} \tag{7}$$

$$\text{RSD}\left[\widetilde{\beta}\right] = \frac{\text{sd}\left[\widetilde{\beta}\right]}{\overline{\beta}}$$

where sd is the standard deviation of the distribution of the estimates and C is the total number of estimates, *i.e.* the number of simulated datasets. A RSD <10% for the fixed effect parameters and <20% for the random effect parameters was considered as precise parameter estimates.

The RBias and the RSD were calculated for each population parameter under each scenario investigated. When all data were used to estimate the parameters the RBias and the RSD were calculated as in (6) and (7) but the estimates of β $\left(\widehat{\beta}\right)$ was used instead of $\widetilde{\beta}$.

RESULTS

The individual estimates of CL, received after fitting the base model to simulated data, was summarized in histograms to show their distributions and overlap for the different scenarios (Fig. 1). The η-shrinkage was calculated for each scenario and the values are presented in Table I (column 5).

The RBias and RSD of the estimated fixed and random effects are presented in Table I (RBias, columns 8–11; RSD columns 12–15) and the bias and precision of the fixed effect estimates of CL are visualized in box plots, showing the bias as the deviation of the median estimate from the true value and the precision as the width of the box and the whiskers (Fig. 2). All population parameters were unbiased under all scenarios. The precision decreased with increasing shrinkage for both the fixed effects and the BSV. For the scenarios with highest shrinkage, *i.e.* scenarios 4 and 8, the precision of the fixed effect of CL for females (CL$_{\text{female}}$) was on the border of the predefined limit (RSD, 10%). For all scenarios with increased shrinkage, *i.e.* scenarios 2–4 and 5–8, the precision of the BSV was lower than the predefined limit (RSD, 30–70%). For all tested scenarios the MI method gave similar bias and precision as when all data were used in the analysis.

The percentage of data sets for which the covariate effect was significant was calculated for each scenario and is reported in the 7th column of Table I.

Table I. Settings for the Simulated Scenarios (Columns 2–4), Calculated Shrinkage (Column 5) and Relative Bias (RBias) [%] and Relative Standard Deviation (RSD) [%] in Population Parameter Estimates (Columns 8–15)

Scenario	Sex difference in CL	Individuals with only 1 observation	Residual error	Shrinkage	Method	Sign. covariate	RBias				RSD			
							CL_{male}	CL_{female}	BSV	ResErr	CL_{male}	CL_{female}	BSV	ResErr
1	17%	0	20%	8.8%	ALL	97%	−0.32	−0.066	−1.7	0.82	3.0	3.8	13	11
					MI	90%	−0.63	0.20	−1.8	0.82	3.5	4.7	14	11
2	17%	0	50%	33%	ALL	76%	−0.23	0.053	−5.6	0.83	4.0	5.2	31	11
					MI	70%	−0.33	0.063	−7.2	0.82	5.0	7.0	32	11
3	17%	2/3	50%	42%	ALL	68%	0.55	0.84	−3.1	−0.40	5.1	5.8	44	17
					MI	66%	0.52	0.95	−4.8	−0.41	5.6	7.9	46	17
4	17%	2/3	70%	54%	ALL	45%	0.87	1.2	−2.3	−0.92	6.5	7.8	69	15
					MI	46%	0.68	1.7	−4.4	−1.0	7.2	10	70	15
5	50%	0	20%	4.5%	ALL	100%	−0.32	−0.064	−1.7	0.82	3.0	3.8	13	11
					MI	100%	−0.69	0.14	0.75	0.82	3.5	4.6	16	11
6	50%	0	50%	21%	ALL	100%	−0.23	0.052	−5.6	0.82	4.0	5.2	31	11
					MI	100%	−0.68	0.46	−2.9	0.82	5.0	6.7	36	11
7	50%	2/3	50%	28%	ALL	100%	0.55	0.83	−3.1	−0.40	5.1	5.8	44	17
					MI	100%	0.19	0.99	1.8	−1.5	6.1	7.8	47	17
8	50%	2/3	70%	40%	ALL	100%	0.87	1.2	−2.2	−0.92	6.5	7.8	69	15
					MI	100%	0.58	1.5	5.2	−2.1	7.8	10	68	15

The between subject variability (BSV) and the residual error (ResErr) were estimated as variances
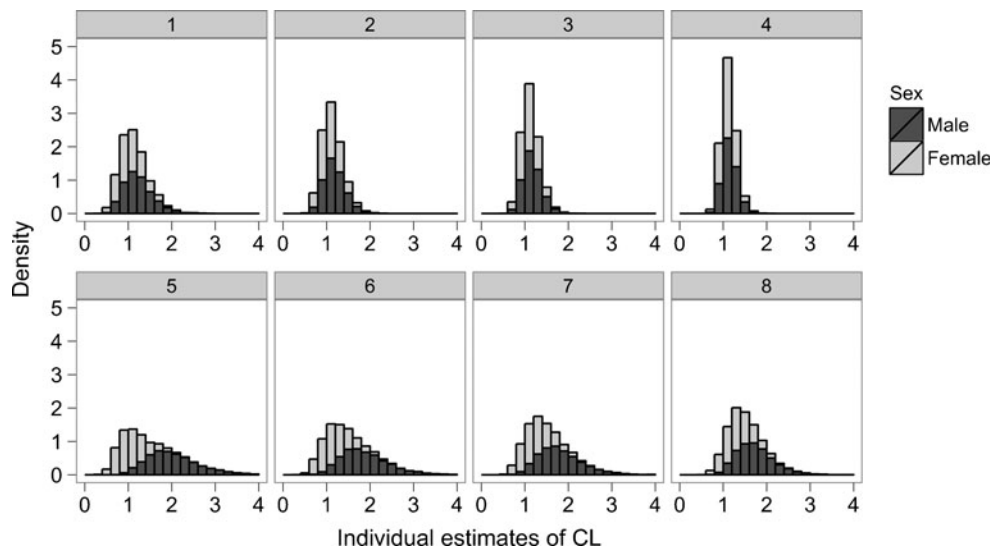
Fig. 1. Distribution of individual CL estimates received after fitting data for 10,000 individuals (6,000 males and 4,000 females) to the base model. The distribution of individual estimates for the females (*light grey*) starts at the upper end of the distribution of individual estimates for the males (*dark grey*). The *panels* present the distribution of the individual estimates for the different scenarios investigated

DISCUSSION

The MI method, first presented by Wu and Wu (16), was successfully implemented in NONMEM and automated using PsN. The method had four steps and in the first step the individual parameters were estimated from a base model (Eq. 2). Wu and Wu suggest that this model should not contain any covariates. However, covariates which are fully observed and which are known *a priori* to be of importance for the PK/PD (*e.g.* body weight in a pediatric study) can be included in the base model to reduce the unexplained

differences between subjects and thus give a clearer signal of the information available about the partly missing covariate. Note that the vector $\widehat{\theta}_i^{(0)}$ is a function of the estimates of θ ($\widehat{\theta}$) and the estimates of η_i ($\widehat{\eta}_i$) and in the present example does not include values of any covariates. For proper imputations it is crucial to include at least as much information as will be used in the final PK/PD model in the regression model and that comprises the covariates used in the base model and the information extracted from the dependent variable in the base model. Since the imputations benefit from inclusion of many descriptive variables, additional covariates
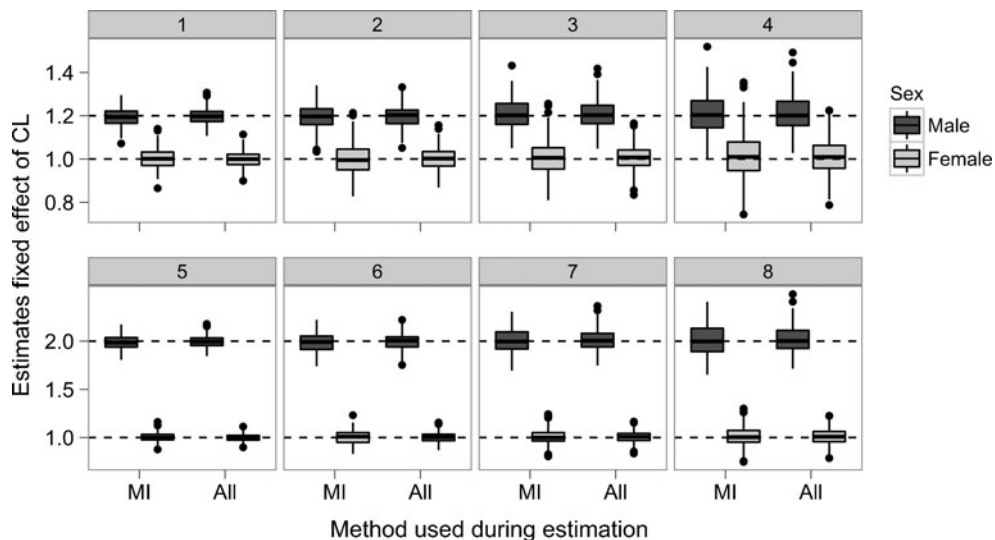

Fig. 2. Box-plots showing bias and precision of the estimated fixed effects of CL. The *panels* present the results for the different scenarios and 200 data sets were simulated and thereafter estimated for each scenario. The covariate sex was missing at random for 50% of the individuals and the multiple imputation method (*MI*) was used to handle the missing data in the estimations, here compared with the results received when no data was missing (*All*)

may be advantageously included in the regression model (17–20). The relationships between the descriptive variables and the partly missing covariates in the regression model (Eq. 3) are linear in the paper by Wu and Wu (16). Other relationships can be tested with the implementation in NONMEM but caution has to be taken against over parameterization of the model. With the current notation in Eq. (3) the residual error model can only have one parameter. However, there is nothing preventing a more advanced error model if there is information in the data to support it. Moreover, a residual error model with one homoscedastic (additive) and one heteroscedastic (proportional) component would enable different types of errors for different parameters. When data have been imputed for the missing covariate values m times, the m completed data sets are analysed and the final parameter estimates are combined according to (4) to obtain one set of final estimates (6,19). The random effects in NONMEM are assumed to be normally distributed and the estimates in Ω and Σ are point estimates of the variances and covariances (24). Therefore it is valid to use the equations in (4) to calculate the combined final estimates also for these parameters.

The implemented method was evaluated for different levels of η-shrinkage. The shrinkage was increased by decreasing the amount of information on the individual level in the simulated data sets, i.e. the residual error was increased and/or the number of samples per individual in the data sets was decreased. The method was also evaluated for small and large covariate effects (17 and 50% difference in CL between males and females). The distributions of individual estimates of CL for males and females overlapped more with increasing η-shrinkage but when the covariate effect was small the overlap was large for lower shrinkage levels as well (Fig. 1, top panel). That explains why the covariate was not found significant in all data sets even when shrinkage was low (Table I, column 7). The MI method gave similar results as when all data were used in the analysis for all scenarios tested, independent on covariate effect and η-shrinkage. When the amount of information on the individual level decreases and the individual estimates shrink towards the fixed effects, the amount of information available in the dependent variable about the partly missing covariate decrease. The MI method, which extracts information about the partly missing covariate from the dependent variable, is hence insensitive to η-shrinkage even though the individual estimates, used in the imputations, suffer from shrinkage.

MI is a flexible method when it comes to advanced relationships between variables and when more than one variable is partly missing at a time. The implementation in PsN makes MI readily accessible to people working in the field of nonlinear mixed effects modelling of clinical data. However, with the method as presently implemented in PsN, it is only possible to impute missing values for one covariate at a time while other software implementations of the MI methodology impute missing data for all partly missing variables at the same time (10–15). It is possible to extend the MI method presented in this paper by adding extra regression models, one for each partly missing covariate, and include all the established relationships in the imputation model. The modelling procedure would not be completely straight forward but no other software than NONMEM would be needed to analyse a data set with partly missing covariate data using MI.

CONCLUSIONS

A MI method for handling of missing covariate data was successfully implemented in NONMEM and automated in PsN. The method gave unbiased parameter estimates independent on level of η-shrinkage and the precision of the estimates are comparable to those received when all data were used in the estimations.

REFERENCES

1. Rubin DB. Inference and missing data. Biometrika. 1976;63 (3):581–92.
2. Orchard T, Woodbury MA, editors. A missing information principle: theory and applications. Proceedings of the sixth Berkeley symposium on mathematical statistics and probability; 1972.
3. Buck SF. A method of estimation of missing values in multivariate data suitable for use with an electronic computer. J R Stat Soc Ser B Methodol. 1960;22(2):302–6.
4. Walsh JE. Computer-feasible method for handling incomplete data in regression analysis. J ACM. 1961;8(2):201–11.
5. Rubin DB, editor. Multiple imputations in sample surveys—a phenomenological Bayesian approach to nonresponse. Proceedings of the survey research methods section: American Statistical Association; 1978.
6. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
7. Goldstein H, Carpenter J, Kenward MG, Levin KA. Multilevel models with multivariate mixed response types. Stat Model Int J. 2009;9(3):173–97.
8. Foulkes AS, Yucel R, Reilly MP. Mixed modeling and multiple imputation for unobservable genotype clusters. Stat Med. 2008;27(15):2784–801.
9. Schafer JL, Yucel RM. Computational strategies for multivariate linear mixed-effects models with missing values. J Comput Graph Stat. 2002;11(2):437–57.
10. Su YS, Gelman A, Hill J, Yajima M. Multiple imputation with diagnostics (mi) in R: opening windows into the black Box. J Stat Softw. 2011;45(2):1–31.
11. Buuren SV, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations. R J Stat Softw. 2011;45(3):1–67.
12. Royston P, White IR. Multiple Imputation by Chained Equations (MICE): implementation in stata. J Stat Softw. 2011;45(4):1–20.
13. Carpenter JR, Goldstein H, Kenward MG. REALCOM-IMPUTE software for multilevel multiple imputation with mixed response types. J Stat Softw. 2011;45(5):1–14.
14. Yuan Y. Multiple imputation using SAS software. J Stat Softw. 2011;45(6):1–25.
15. Honaker J, King G, Blackwell M. Amelia II: a program for missing data. J Stat Softw. 2011;45(7):1–47.
16. Wu H, Wu L. A multiple imputation method for missing covariates in non-linear mixed-effects models with application to HIV dynamics. Stat Med. 2001;20:1755–69.

17. Meng XL. Multiple-imputation inferences with uncongenial sources of input. Statistical Science. 1994;538–58.

18. Rubin DB. Multiple imputation after 18+ years. Journal of the American Statistical Association. 1996;473–89.

19. Schafer JL. Analysis of incomplete multivariate data: Chapman & Hall; 1997.

20. Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychol Methods. 2001;6(4):330.

21. Maitre P, Bührer M, Thomson D, Stanski D. A three-step approach combining Bayesian regression and NONMEM population analysis: application to midazolam. J Pharmacokinet Pharmacodyn. 1991;19(4):377–84.

22. Mandema JW, Verotta D, Sheiner LB. Building population pharmacokinetic-pharmacodynamic models. I. Models for covariate effects. J Pharmacokinet Pharmacodyn. 1992;20(5): 511–28.

23. Sheiner LB, Beal SL. Bayesian individualization of pharmacokinetics: simple implementation and comparison with non-Bayesian methods. J Pharm Sci. 1982;71(12):1344–8.

24. Beal S, Sheiner LB, Boeckmann A, Bauer RJ. NONMEM User's Guides. (1989–2009). ICON Development Solutions, Ellicott City, MD, USA; 2009.

25. Xu XS, Yuan M, Karlsson MO, Dunne A, Nandy P, Vermeulen A. Shrinkage in nonlinear mixed-effects population models: quantification, influencing factors, and impact. AAPS J. 2012;14(4):927–36.

26. Savic RM, Karlsson MO. Importance of shrinkage in empirical Bayes estimates for diagnostics: problems and solutions. AAPS J. 2009;11(3):558–69.

27. Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN)—a Perl module for NONMEM related programming. Comput Methods Prog Biomed. 2004;75(2):85–94. Epub 2004/06/24.

28. Lindbom L, Pihlgren P, Jonsson N. PsN-toolkit—a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. Comput Methods Prog Biomed. 2005;79(3):241–57.

29. Tunblad K, Lindbom L, McFadyen L, Jonsson EN, Marshall S, Karlsson MO. The use of clinical irrelevance criteria in covariate model building with application to dofetilide pharmacokinetic data. J Pharmacokinet Pharmacodyn. 2008;35(5):503–26.