

# Analysis of non-ignorable missing and left-censored longitudinal data using a weighted random effects tobit model

Abdus Sattar,<sup>a\*†</sup> Lisa A. Weissfeld<sup>b</sup> and Geert Molenberghs<sup>c</sup>

In a longitudinal study with response data collected during a hospital stay, observations may be missing because of the subject's discharge from the hospital prior to completion of the study or the death of the subject, resulting in non-ignorable missing data. In addition to non-ignorable missingness, there is left-censoring in the response measurements because of the inherent limit of detection. For analyzing non-ignorable missing and left-censored longitudinal data, we have proposed to extend the theory of random effects tobit regression model to weighted random effects tobit regression model. The weights are computed on the basis of inverse probability weighted augmented methodology. An extensive simulation study was performed to compare the performance of the proposed model with a number of competitive models. The simulation study shows that the estimates are consistent and that the root mean square errors of the estimates are minimal for the use of augmented inverse probability weights in the random effects tobit model. The proposed method is also applied to the non-ignorable missing and left-censored interleukin-6 biomarker data obtained from the Genetic and Inflammatory Markers of Sepsis study. Copyright © 2011 John Wiley & Sons, Ltd.

**Keywords:** longitudinal data; non-ignorable missing; left-censoring; inverse probability weighting; pseudo-likelihood method

## 1. Introduction

Missing data are a persistent problem in longitudinal studies, presenting challenges at the analysis stage and resulting in a need for methodology to address these issues. The data may be missing because of subject dropout or death, failure to collect a subset of the data at follow-up for administrative reasons, or censoring. This loss of data can result in biased estimates and a reduction in precision. In addition, the relationship between the outcome and predictors may vary depending on the reason for missingness, and failure to account for this in an analysis can affect the results.

Many methods exist for the handling of missing data in longitudinal studies when the data are missing at random (MAR) and missing not at random (MNAR) including those based on imputation [1] and those based on weighting [2]. Using the method of weighting, greater weight is placed on those observations that are less likely to be observed. The attraction of both imputation and weighting is that standard methods can be used for analysis [3]. There are also likelihood-based methods, including selection, pattern mixture, and shared parameter models, for analyzing MNAR or non-ignorable missing longitudinal data [4, 5]. In these likelihood-based methods, joint models for the response and missingness indicator variables are formulated for a valid statistical inference. By combining the features of the selection model and the pattern-mixture model, Yuan and Little [6] developed a new class of models, called mixed-effects hybrid models (MEHM). In this new MEHM, the joint likelihood of the response and missingness indicator has been factored into three parts: the marginal distribution of random effects, the conditional distribution of the dropout process given the random effects, and the conditional distribution of the

<sup>a</sup>Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA

<sup>b</sup>Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA, USA

<sup>c</sup>I-BioStat, Universiteit Hasselt, B-3590, Diepenbeek, Belgium

\*Correspondence to: Abdus Sattar, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA.

†E-mail: sattar@case.edu

response given the random effects and dropout process. This class of model is also a generalization of shared parameter models and computationally easy to implement. The detail of the formulation of the likelihood function and the parameter estimation associated with the above likelihood-based MEHM can be found in Fitzmaurice *et al.* [7] and Yuan and Little [6].

An alternative to the likelihood-based methods is the weighting-based methods that do not require joint modeling of the longitudinal response and missingness indicator vector. The weighting-based methods require two separate models, one for the response variable and the other for the response indicator variable. Weights are computed using the response indicator variable models at the first stage. Then these estimated weights are used in the response model to adjust for the non-response. Weighting-based methods have roots in the Horvitz–Thompson estimators in survey sampling where inverse probability weighting (IPW) has been used for the differential sampling rate problem to account for the fact that the data are not obtained from a random sample [8]. The idea of the IPW method is that if the probability of selecting a unit is  $\pi_i$ , then  $\pi_i^{-1}$  units in the population should be used in the estimation. Robins and Rotnitzky [9] introduced the ‘augmented’ IPW (IPWA) method for an incompletely observed response profile in both non-parametric and semi-parametric models. This IPWA estimator has a close connection with the difference and regression estimators of the population total and means in the survey sampling literature. The distinguishing feature of the IPWA estimator is that it provides an unbiased estimate when the missingness is due to the design. An efficient IPWA estimator can be obtained by an ingenious choice of the augmentation term compared with the non-augmented Horvitz–Thompson-type IPW estimators [2]. For the case of non-ignorable missing longitudinal data, the IPWA method has been used in these settings.

Because of inadequate instrument sensitivity, a large percentage of response measurements are often subject to a limit of detection (LOD) in many longitudinal studies. When the observed data fall below or above the LOD, these data will be referred to as censored data. Although a censored data value can be viewed as a type of missing data, the data are not actually missing because some information is available owing to the recording of the censoring limit. Censored data also satisfy the definition of non-ignorable missingness because the censoring is related to the unobserved values [10]. Because left-censored data arise in a variety of applications, there are many methods available to account for left-censoring in the outcome, with the tobit regression model being one of the first models developed for this problem [11]. The tobit regression model has been extended for use with longitudinal data by introducing random effects into the model [12]. Other approaches include imputation of the quantification limit, using half of the lower limit of detection, and using random imputation procedures. Hughes [13] proposed a more efficient Monte Carlo expectation–maximization (MCEM) algorithm for analyzing censored longitudinal data using a linear mixed model. This approach uses Gibbs sampling to implement a Monte Carlo expectation step in the expectation–maximization algorithm, and the resulting estimates are likely to give reasonable results in a wide range of situations. A more flexible approach has also been developed by formulating and maximizing the likelihood function for the observed and censored data directly [14]. In this approach, the contribution of the censored data to the likelihood function is obtained by using the multivariate conditional distribution of the censored response given the observed response. Censored data, as well as non-ignorable missing data, have been analyzed directly by maximizing the joint likelihood of the repeated measurements and time-to-dropouts data [15]. This approach avoided the MCEM algorithm, and the standard errors (SE) of the estimates were computed using the information matrix. Longitudinal response measurements with censoring and non-ignorable missingness have also been analyzed by the shared parameter approach in the joint modeling framework [16]. These two joint modeling approaches differ primarily in the choices of the dropout process models.

The goal of this work was to address the problem of non-ignorable missing longitudinal data when the outcome is subject to left-censoring. We are proposing to extend the theory of the random effects tobit regression model [12] to develop a weighted random effects tobit (WRT) model for analyzing non-ignorable missing and left-censored longitudinal data. We will compare the performance of the WRT model with the random effects tobit (RT) model [12] as well as several versions of the weighted linear mixed model [17]. In this setting, the weighted linear mixed models (WMM) will be fit by replacing the censored values with either half of the detection limit or a randomly imputed value. In sample survey theory, the weights have been assumed to be fixed and known. Here we compute the weights from the observed data, and hence we will take into account their sampling variability in the inference. We will consider the weights as nuisance parameters in the WRT model and base our inference on pseudo-likelihood (PL) theory to account for the uncertainty associated with the estimation of a large number of nuisance parameters [18].

In Section 2, we will describe the random effects tobit model and our proposed weighted version of this model for analyzing non-ignorable missing and left-censored longitudinal data. In Section 3, we will demonstrate the performance of the proposed model using a simulation study. In Section 4, we will present an analysis of the interleukin-6(IL6) biomarker data from the Genetic and Inflammatory Markers of Sepsis (GenIMS) study and will also perform a sensitivity analysis of the presented results. In the final section, we will offer a discussion on the findings.

## 2. Notation and weighted random effects tobit model formulation

Let  $Y_{ij}$  denote the measurement of a response variable from the  $i$ th subject at the  $j$ th wave of measurements at time  $t_{ij}$ ,  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, n_i$  and  $X_{ij} = (X_{ij1}, X_{ij2}, \dots, X_{ijp})^T$  denote a  $p \times 1$  vector of covariates associated with  $Y_{ij}$ . In vector notation,  $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{in_i})^T$  denotes the  $n_i$ -dimensional vector of the response measurements, and  $X_i$  denotes the  $n_i \times p$  matrix of covariates for the  $i$ th subject. Suppose that the observed and missing values of  $Y_{ij}$  are denoted by  $Y_{ij}^o$  and  $Y_{ij}^m$ , respectively. Define the missingness indicator vector  $R_i = (R_{i1}, R_{i2}, \dots, R_{in_i})^T$  where  $R_{ij} = 1$  if  $Y_{ij}$  is observed, and 0 otherwise. Following Rotnizky *et al.* [19], let  $\bar{R}_{ij} = (R_{i1}, R_{i2}, \dots, R_{i,j-1})^T$  and  $\bar{Y}_{ij}^o = (Y_{i1}^o, Y_{i2}^o, \dots, Y_{i,j-1}^o)^T$ , and assume that  $P(R_{ij} | \bar{R}_{ij} = \mathbf{1}, \bar{Y}_{ij}^o, X_{ij})$  is bounded away from zero with probability 1 and that  $\mathbf{1}$  is a response vector of 1s up to but not including time  $j$ . We further assume that  $Y_i$  and  $Y_{i'} (i \neq i')$  are independent and that the covariate vector  $X_{ij}$  is fully observed. Let  $f(\cdot)$  and  $F(\cdot)$  denote the probability density function and cumulative density function, respectively.

Suppose we observe  $Y_{ij} = Y_{ij}^o$  only if  $Y_{ij}^o > c$  (a constant) and  $Y_{ij} = c$  if  $Y_{ij}^o < c$ . In this scenario, we have left-censored observations because we do not observe any  $Y_{ij}$  that is less than  $c$ . Let  $d_{ij} = 1$  if  $Y_{ij}$  is observed and  $d_{ij} = 0$  if  $Y_{ij}$  is censored. The left-censored longitudinal data can be analyzed utilizing the following tobit model [11]:

$$Y_{ij}^o = X_{ij}^T \beta + \varepsilon_{ij} \quad \text{if } d_{ij} = 1, \quad (1)$$

where  $\beta$  is a  $p \times 1$  parameter vector and  $\varepsilon_{ij}$  is the error component that is assumed to be independent and identically distributed as  $N(0, \sigma^2)$ . To estimate the parameters in model (1), we can use the maximum likelihood procedure. Contributions of the uncensored and censored observations to the likelihood function are  $f(Y_{ij}^o | X_{ij}; \beta, \sigma^2)$  and  $F(c | X_{ij}; \beta, \sigma^2) = P(Y_{ij}^o < c | X_{ij}; \beta, \sigma^2)$ , respectively. From these probability specifications, the likelihood function for the parameters associated with model (1) is given by the following:

$$L(\beta, \sigma^2) = \prod_{i=1}^N \prod_{j=1}^{n_i} [f(Y_{ij}^o | X_{ij}; \beta, \sigma^2)]^{d_{ij}} [F(c | X_{ij}; \beta, \sigma^2)]^{1-d_{ij}} \quad (2)$$

Epstein *et al.* [12] have developed a tobit model in the linear mixed model (LMM) frame work. We will refer to this new model as the RT model. A LMM for the  $i$ th subject can be written in the following form [17]:

$$Y_i = X_i \beta + Z_i b_i + \varepsilon_i, \quad (3)$$

where the vector  $Y_i$  is distributed as multivariate normal with the following specification:

$$Y_i | b_i \sim N(X_i \beta + Z_i b_i, \Sigma_i), \quad b_i \sim N(0, D), \quad (4)$$

where,  $X_i$  is the  $n_i \times p$  design matrix for the fixed effects,  $\beta$ ;  $Z_i$  is the  $n_i \times q$  design matrix for the random effects,  $b_i$ ; and  $D = \text{Cov}(b_i)$  and  $\Sigma_i = \text{Cov}(\varepsilon_i)$  are the covariance matrices of the random effects and errors, respectively. If there is no left-censoring in the measurements, then it follows from (3) that the marginal distribution of  $Y_i$  is normal with mean  $X_i \beta$  and covariance matrix  $V_i = \Sigma_i + Z_i D Z_i^T$ . So the inferences for the fixed parameters  $\beta$  can be based on the log-likelihood function,  $\log L(\beta) \propto -0.5 \log |V_i| - 0.5(Y_i - X_i \beta)^T V_i^{-1} (Y_i - X_i \beta)$ . For the left-censored response measurements, the likelihood function for the RT model [12] can be written as

$$L(\beta, \eta) = \prod_{i=1}^N \int \left\{ \prod_{j=1}^{n_i} [f(Y_{ij}^o | X_{ij}, Z_{ij}; b_i, \beta, \eta)]^{d_{ij}} [F(c | X_{ij}, Z_{ij}; b_i, \beta, \eta)]^{1-d_{ij}} \right\} \times f(b_i | D) db_i, \quad (5)$$

where  $\eta$  denotes the vector of all parameters in the variance–covariance matrices  $D$  and  $\Sigma_i$ . This likelihood function does not have an analytic or closed-form expression. Estimation of  $\beta$  in the presence of the nuisance parameters  $\eta$  can be based on the pseudo-maximum likelihood method [18]. When applying the pseudo-maximum likelihood method, the likelihood function is maximized for the parameters of interest, and all other parameters are treated as nuisance parameters. These nuisance parameters are replaced by their consistent estimates in the likelihood function. Treating  $\eta$  as a nuisance parameter vector, the pseudo-log-likelihood and the corresponding score function for the parameter vector  $\beta$  is as follows:

$$l(\beta, \hat{\eta}) = l(\beta, \eta)|_{\eta=\hat{\eta}} \quad (6)$$

$$S(\beta, \hat{\eta}) = \frac{\partial}{\partial \beta} l(\beta, \eta)|_{\eta=\hat{\eta}}. \quad (7)$$

Equations (6) and (7) define the log-likelihood function and the score function corresponding to the RT model, respectively. If the missing data are ignorable, then we can base the inferences on the RT model's pseudo-log-likelihood function (6) and its score Equation (7). For left-censored longitudinal data subject to non-ignorable missingness, the likelihood function is complex [10]. It requires integrals not only for the censored observations but also for the random effects and the missing data distributions. These integrals do not have a closed form, further adding to the computational complexity of the problem. Hence, these often need numerical approximations with the uncertainty of the convergence in the estimation process. To simplify this likelihood, we implement the pseudo-likelihood method with weighting to account for missing data [20]. We incorporate these weights by multiplying (dot product) the pseudo-log-likelihood function (Equation 6) and the score function (Equation 7) of the RT model by weights. Now, the weighted pseudo-log-likelihood function and weighted score function of the RT model can be defined as

$$l(\beta, \hat{\eta}, \hat{\omega}) = \sum_{i=1}^N \frac{I(R_i = 1)}{\pi_i} \cdot l_i(\beta, \hat{\eta}) \quad (8)$$

$$S(\beta, \hat{\eta}, \hat{\omega}) = \sum_{i=1}^N \frac{I(R_i = 1)}{\pi_i} \cdot \frac{\partial}{\partial \beta} l_i(\beta, \hat{\eta}), \quad (9)$$

where  $\omega = \{\omega_i = \pi_i^{-1}, i = 1, \dots, N\}$ ;  $\pi_i$  is the probability (vector) of observing the response from the  $i$ th subject and assuming that these probabilities are bounded away from zero. We assume that the estimate of  $\pi_i$  will be consistent. The probability of observing a response is modeled with a logit model; however, other models can be used to estimate these probabilities. We will fit the logistic regression model with the covariate vector  $X_{ij}$ , the observed response (or outcome) vector prior to time  $t_j$ ,  $\bar{Y}_{ij}^o$ , and the unobserved response  $Y_{ij}^m$  at time  $t_j$  as described by Rotnitzky *et al.* [19, p.1325]. Thus, we can write the logistic regression model as

$$\text{logit}(\pi_{ij}) = \lambda_0 + \lambda_1^T X_{ij} + \lambda_2^T \bar{Y}_{ij}^o + \lambda_3 Y_{ij}^m. \quad (10)$$

Because  $Y_{ij}^m$  is not observable at time  $t_j$ , we will adopt the IPWA method to compute  $\pi_{ij}$ . Initially, we will treat  $\lambda_3$ , the augmented value, as a known constant and estimate the other coefficients by solving [19]

$$\sum_i \sum_{r \neq 1} \left\{ I(R_i = r) - \frac{I(R_i = 1)}{\pi_i(1; \lambda)} \pi_i(r; \lambda) \right\} \gamma_r(X_i, Y_i) = 0 \quad (11)$$

In Equation (11), the set  $\{r = (r_1, \dots, r_{n_i}) : r_j = 0, 1\}$  and  $\gamma_r(\cdot)$  is an arbitrary  $(p-1)$  vector function of the observed data, where  $p-1$  is the number of unknown parameters to be estimated. Rotnitzky *et al.* [19] suggest considering a series of plausible fixed values for the augmented term,  $\lambda_3$ , followed by an evaluation of the sensitivity of the inference for the parameters of interest. We will perform the sensitivity analysis in Section 4.2 and provide more details on the augmented value selection and evaluation of the inferences.

The inference for the pseudo-likelihood estimates obtained from the score Equation (9) is based on asymptotic PL theory. The asymptotic SE of the estimates accounts for the extra variability due to the

use of the estimated weights in the estimation process [18]. We will obtain the asymptotic covariance matrix for the parameter of interest by using the following formula [20]:

$$\text{Var}(\hat{\beta}) \cong A_{11}^{-1} + A_{11}^{-1} \tilde{V} A_{11}^{-1} \quad (12)$$

where

$$A_{11}^{-1} = - \sum_{i=1}^N R_i \hat{\omega}_i \cdot \frac{\partial}{\partial \beta \partial \beta^T} l_i(\beta, \hat{\eta})|_{\beta=\hat{\beta}}. \quad (13)$$

Letting  $\tilde{\xi}_i = \frac{\partial}{\partial \beta} l(\beta, \eta)$  and  $\tilde{\xi}^{(g)} = \frac{1}{n_g} \sum_{i \in D_g} \tilde{\xi}_i$ ,  $\tilde{V}$  can be defined as

$$\tilde{V} = \sum_{i=1}^N \omega_i (\omega_i - 1) \sum_{i \in D_g} \left( \tilde{\xi}_i - \tilde{\xi}^{(g)} \right) \left( \tilde{\xi}_i - \tilde{\xi}^{(g)} \right)^T, \quad (14)$$

where  $D_g$  is the set of observations in the  $g$ th stratum or group,  $g = 1, 2, \dots, K$ . We have derived the variance formula (12) and its component  $\tilde{V}$  using the conditional variance of the scores conditioning on  $Y_i$ ,  $X_i$ , and  $Z_i$ , so that  $\text{Var}(S_i) = \text{Var}\{E(S_i|Y_i, X_i, Z_i)\} + E\{\text{Var}(S_i|Y_i, X_i, Z_i)\}$ . On the basis of the formulation by Lawless *et al.* [20] of the weighted pseudo-likelihood theory, we have defined  $g$  strata or groups of subjects (e.g., subjects who completed the study, the dropouts from the study, and those who died during the study, etc.) in our analysis. Then we have computed the probability of missing an observation for each subject and modified the contribution of an observation to the likelihood function through the use of weights. We then compute the adjusted SEs of the estimates through the variance formulas (12)–(14).

Equations (8) and (9) are the log-likelihood function and score equation of the proposed WRT model, respectively. We will compute the asymptotic SE of the WRT model's parameter estimate using formulas (12)–(14). We will compare this SE with the SE based on the observed Fisher information matrix and the sandwich estimator. We compute the SE based on the observed Fisher information from  $[\mathcal{I}(\hat{\beta})]^{-1}$ , where  $\mathcal{I}(\hat{\beta}) = -\partial S(\cdot)/\partial \beta$ , and  $S(\cdot)$  represents score functions as defined previously. We compute the sandwich estimator by the following formula:

$$[\mathcal{I}(\hat{\beta})]^{-1} \left( \sum_{i=1}^N [R_i \hat{\omega}_i \cdot \{\partial l_i(\beta, \hat{\eta})/\partial \beta\}] [R_i \hat{\omega}_i \cdot \{\partial l_i(\beta, \hat{\eta})/\partial \beta\}]^T \right) [\mathcal{I}(\hat{\beta})]^{-1}.$$

For comparison, we will fit a WMM to the non-ignorable missing and left-censored longitudinal data. In the WMM, we will use the IPWA method to address the non-ignorable missingness and will replace the censored values with half of the detection limit or a randomly imputed value. The score equation for the parameters of interest from the WMM will be as follows,

$$\frac{\partial \log L(\beta)}{\partial \beta} = -X_i^T V_\omega^{-1} X_i \beta + X_i^T V_\omega^{-1} Y_i, \quad (15)$$

where  $\text{Cov}(Y) = V_\omega = ZGZ^T + LRL^T$  with  $L = \text{diag}(\omega^{-1/2})$ . We will estimate the parameters of the WMM by the score Equation (15) and will compute the SEs of the estimated parameters by the variance formula in (12)–(14).

We will compare the performance of the proposed WRT model with a number of different models. We denote each of these models with a subscript on both the left and right sides. The subscripts on the left side are i, a, and r denoting the estimation technique used to obtain the SE. In this case i, a, and r denote the estimate based on the Fisher information matrix, the asymptotic SE from the PL theory, and the robust SE, respectively. The subscripts on the right side are as follows: 1, ipw, ipwa, mis, and large denoting the type of weights used across the wave of measurements; y-mis denotes the response model is mis-specified; h and r denote that the left-censored response values are replaced by half of the detection limit and a randomly imputed value, respectively. In this case, 1, ipw, ipwa, mis, and large denote the use of a single weight, IPW, IPWA, mis-specified weights, and large weights in the model, respectively.



The models that will be considered are the following

- (1)  $i$ MM: denotes a mixed-effects model with the SEs of the estimates computed using the Fisher information matrix. This mixed model is fitted using the complete simulated data, that is, there is no censoring, and there are no missing observations.
- (2)  $i$ RT: denotes a random effects tobit model with the SEs of the estimates computed using the Fisher information matrix. This model assumes that the missing data mechanism is MAR.
- (3)  $i$ WRT<sub>1</sub>: denotes a weighted random effects tobit model with the SEs of the estimates computed using the Fisher information. Single weights are computed using the baseline covariates and the first day's measurements.
- (4)  $a$ WRT<sub>1</sub>: denotes a weighted random effects tobit model with the SEs of the estimates computed using the asymptotic PL theory. Single weights are computed using the baseline covariates and the first day's measurements.
- (5)  $a$ WRT<sub>ipw</sub>: denotes an IPW random effects tobit model with the SEs of the estimates computed using the asymptotic PL theory.
- (6)  $a$ WRT<sub>ipwa</sub>: denotes an IPWA random effects tobit model with the SEs of the estimates computed using the asymptotic PL theory.
- (7)  $a$ WRT<sub>mis</sub>: denotes a weighted (IPWA) random effects tobit model with the SEs of the estimates computed using the asymptotic PL theory. In this model, a small perturbation of IPWA has been applied to 50% of randomly chosen weights, and hence, the weights are mis-specified.
- (8)  $r$ WRT<sub>mis</sub>: denotes a weighted (IPWA) random effects tobit model with the SEs of the estimates computed using the sandwich estimator. In this model, a small perturbation of IPWA has been applied to 50% of randomly chosen weights, and hence, the weights are mis-specified.
- (9)  $a$ WRT<sub>large</sub>: denotes a weighted (IPWA) random effects tobit model with the SEs of the estimates computed using the asymptotic PL theory. In this model, a large perturbation of IPWA has been applied to 10% of randomly chosen weights, and hence, some weights are large and mis-specified.
- (10)  $a$ WRT<sub>y-mis</sub>: denotes a weighted (IPWA) random effects tobit model with the SEs of the estimates computed using the asymptotic PL theory. In this model, the functional form of the WRT model has been mis-specified.
- (11)  $a$ WMM<sub>h</sub>: denotes a weighted (IPWA) linear mixed model with the SEs of the estimates computed using the asymptotic PL theory. In this model, the censored values are replaced by half of the detection limit.
- (12)  $a$ WMM<sub>r</sub>: denotes a weighted (IPWA) linear mixed model with the SEs of the estimates computed using the asymptotic PL theory. In this model, the censored values are replaced by randomly imputed values.

### 3. Simulation study for non-ignorable missing and left-censored longitudinal data

To compare the performance of the proposed WRT model with the RT model and WMM, we conducted a simulation study that was designed to address the following issues: the handling of the missing data through the estimation of the weights used in the model, the definition of the censored outcome, the estimation of the variance of the estimators and the sensitivity of the model to mis-specification of the weights, and functional form of the response variable. We used several different approaches to compare methods for the estimation of the parameters and to examine the sensitivity of the estimates to several types of model mis-specifications.

We compared the  $i$ WRT<sub>1</sub>,  $a$ WRT<sub>1</sub>,  $a$ WRT<sub>ipw</sub>,  $a$ WRT<sub>ipwa</sub>,  $a$ WRT<sub>mis</sub>,  $r$ WRT<sub>mis</sub>,  $r$ WRT<sub>large</sub>, and  $r$ WRT<sub>y-mis</sub> models, with the left subscripts  $i$ ,  $a$ , and  $r$  indicating SEs based on information matrix, asymptotic pseudo-likelihood theory, and the sandwich estimator, respectively. The impact of the estimation of the variance was examined via the models  $i$ WRT<sub>1</sub>,  $a$ WRT<sub>1</sub>,  $a$ WRT<sub>mis</sub>, and  $r$ WRT<sub>mis</sub>, which are weighted random tobit models with the subscript  $i$  denoting the information matrix,  $a$  denoting the asymptotic variance based on the PL theory, and  $r$  denoting sandwich estimator. The right subscripts of 1 indicate that only one (or a constant) weight is used across the wave of measurements;  $ipw$  and  $ipwa$  indicate that the weights are estimated using IPW and IPWA methodology;  $mis$  and  $large$  indicate that the weights are mis-specified and large, respectively; and  $y-mis$  indicates that the functional form of the response variable model is mis-specified. Mis-specified IPWA weights are generated from a Bernoulli distribution with the parameter set equal to 0.5, resulting in 50% of the IPWA weights being perturbed. The

perturbed IPWA weights are generated by adding  $\sqrt{1/u}$  to the weight resulting in a range of 1.0 to 1.41 for the perturbations. Similarly, *large* weights are constructed by adding  $\sqrt{1/u} (> 100)$  to 10% of IPWA weights where  $u < 0.01$  is a random number drawn from uniform(0, 1) distribution. Thus, this scenario will examine the consequence of violation of missingness probability assumption, not away from zero. In the mis-specified response model, the observed covariate  $X$  is (defined in the model (16)) replaced by  $\exp(X/2)$  to assess the impact of mis-specification of the functional form of the response variable [21]. We are not examining two other possible scenarios—the case where both (the response and missingness probability) models are correct and the case where both models are incorrect. It is known that when both models are correct, then the inference is an expected one. Also, it is known that when both models are incorrect, the estimates are severely biased and lead to an incorrect inference [22]. Finally, the impact of ignoring the left-censoring of the outcome was examined by replacing censored values either by half of the lower limit of detection or by a randomly imputed value with the models  ${}_a\text{WMM}_h$  and  ${}_a\text{WMM}_r$ , respectively. These two models will be considered as mis-specified models in the sense that censored values are replaced by ad hoc response values, and a standard mixed-effects model is then fitted to the data.

We generated outcome data from the following basic linear mixed model,

$$Y_{ij}|X_i, Z_{ij}, b_{0i} \sim N(\beta_0 + \beta_1 X_i + \beta_2 Z_{ij} + b_{0i}, \sigma^2), \quad (16)$$

where,  $X_i$  represents a fixed effects (covariate),  $Z_{ij}$  takes values  $1, \dots, 7$ , (time in the study),  $b_{0i} \sim N(0, 1)$  is the random intercept, and  $\sigma^2 = 2$ . The covariate  $X_i \sim N(\zeta_g, 0.3627^2)$ , where  $g$  represents strata (e.g., strata can be defined for the subjects who completed the study (0), dropped out from the study (1), or died during the study period (2)), and we set  $\zeta_0 = 3.74$ ,  $\zeta_1 = 2.74$ , and  $\zeta_2 = 4.74$ . The true values of the parameters were set to  $\beta_0 = 1.5$ ,  $\beta_1 = 0.5$ , and  $\beta_2 = -0.5$ . This random intercept model (16) leads to compound symmetry for the variance–covariance structure of the generated outcome data. The outcome variable was then left-censored on the basis of rates of 25% and 40% for the simulation studies. Following the method in Demirtas and Hedeker [23], we created the non-ignorable missing data using logistic regression models,  $\text{logit}(\pi_{ij}) = \lambda_0 + \lambda_1 Y_{ij}$  and the choice of (i)  $\lambda^T = (\lambda_0, \lambda_1) = (-0.5, 1.0)$  resulting in approximately 25% missing data or (ii)  $\lambda^T = (\lambda_0, \lambda_1) = (1.25, 3.5)$  resulting in approximately 50% missing data. Now on the basis of the methods in Section 2, we have computed the missingness probability by the following logit model:  $\text{logit}(\pi_{ij}) = \lambda_0 + \lambda_1 X_i + \lambda_2^T \bar{Y}_{ij}^o + \lambda_3 Y_{ij}^m$ .

All simulations were run with three different sample sizes of 250, 500, and 1000 subjects and a follow-up period of 7 days. The results presented are based on 500 simulations (iterations) and three different scenarios (or design). The first scenario, S1, is constructed so that 60% of the subjects have complete observations, that is, data are present for all 7 days, and the remaining 40% of the subjects have missing data because of dropout. For the second scenario, S2, 60% of the subjects have complete data, 30% of the subjects have missing data because of dropout, and the remaining 10% have missing data because of death. Scenario 3, S3, consists of 30% of the subjects with complete data, 60% of the subjects with missing data because of dropout, and the remaining 10% missing because of death. In terms of the dropout categories, this third scenario resembles the GenIMS data set that motivated this work. Results associated with the covariate  $X$  are presented in the tables that follow, and the results associated with the intercept and the coefficient of the  $Z$  (time) are similar to those for the covariate  $X$  and have been omitted from the presentation.

Table I presents the results obtained from scenario S1 where 60% of the subjects have complete observations. The first three columns in the table contain the percentages of censored data, non-ignorable missingness, and the fitted model, respectively. Results are presented for sample sizes of 250 and 1000 subjects with 25% and 40% of the outcomes subject to censoring. Examining Table I, we observe that the bias of the estimates depends on the model and weighting scheme used in the analysis. Clearly, the RT model underestimates the regression parameters. The WRT models that use constant weights across the course of measurements also produce a negative bias. Estimates from the two mixed models are also biased. The WRT model estimates that are based on IPWA method are unbiased. As expected from the theory for small sample sizes, the SE obtained from the asymptotic PL theory is slightly larger than the SE based on the Fisher information matrix. In most cases, the SE obtained from the asymptotic PL theory is slightly smaller than the SE based on the sandwich estimator. The SE of the estimate decreases with an increase in the sample size, and for large sample sizes they are approximately equal for all estimators, except for the mis-specified response model. Because of the rescaling of covariate ( $\exp(X/2)$ ), the mis-specified response model produces severely biased estimates with the SE of the estimate away from

**Table I.** Analysis of simulated non-ignorable missing and left-censored longitudinal data considering scenario 1: 40% subjects were dropouts from the study. Only results related to the covariate  $X$  are presented in the table, and the true value of the coefficient is 0.5.

Data feature and model			$N = 250$				$N = 1000$			
Censored	Missing	Model	Bias	SE	RMSE	CP	Bias	SE	RMSE	CP
No	No	$i$ MM	0.01	0.11	0.11	0.92	0.00	0.06	0.06	0.95
25%	25%	$i$ RT	−0.15	0.12	0.19	0.76	−0.15	0.06	0.16	0.29
		$i$ WRT <sub>1</sub>	−0.11	0.12	0.16	0.79	−0.12	0.06	0.14	0.46
		$a$ WRT <sub>1</sub>	−0.11	0.13	0.17	0.79	−0.12	0.06	0.14	0.46
		$a$ WRT <sub>ipw</sub>	0.05	0.13	0.14	0.87	0.04	0.06	0.07	0.83
		$a$ WRT <sub>ipwa</sub>	0.01	0.13	0.13	0.92	0.00	0.06	0.06	0.94
		$a$ WRT <sub>mis</sub>	0.00	0.13	0.13	0.91	−0.01	0.06	0.06	0.91
		$r$ WRT <sub>mis</sub>	0.00	0.13	0.13	0.92	−0.01	0.07	0.07	0.93
		$a$ WRT <sub>large</sub>	0.18	0.13	0.23	0.62	0.19	0.06	0.20	0.29
		$a$ WRT <sub>y-mis</sub>	−0.32	0.05	0.32	0.00	−0.32	0.02	0.32	0.00
		$a$ WMM <sub>h</sub>	−0.09	0.12	0.15	0.84	−0.10	0.06	0.12	0.53
		$a$ WMM <sub>r</sub>	−0.07	0.14	0.16	0.88	−0.08	0.06	0.11	0.70
40%	25%	$i$ RT	−0.15	0.12	0.19	0.72	−0.15	0.06	0.16	0.32
		$i$ WRT <sub>1</sub>	−0.07	0.12	0.14	0.88	−0.07	0.06	0.09	0.80
		$a$ WRT <sub>1</sub>	−0.07	0.14	0.15	0.88	−0.07	0.06	0.09	0.80
		$a$ WRT <sub>ipw</sub>	0.09	0.14	0.17	0.86	0.09	0.06	0.11	0.63
		$a$ WRT <sub>ipwa</sub>	0.00	0.13	0.13	0.93	0.01	0.06	0.06	0.94
		$a$ WRT <sub>mis</sub>	−0.01	0.14	0.14	0.91	0.00	0.06	0.06	0.92
		$r$ WRT <sub>mis</sub>	−0.01	0.13	0.13	0.92	0.00	0.07	0.07	0.94
		$a$ WRT <sub>large</sub>	0.24	0.15	0.28	0.52	0.25	0.07	0.26	0.16
		$a$ WRT <sub>y-mis</sub>	−0.30	0.05	0.30	0.00	−0.30	0.02	0.31	0.00
		$a$ WMM <sub>h</sub>	−0.11	0.12	0.16	0.81	−0.10	0.06	0.12	0.58
		$a$ WMM <sub>r</sub>	−0.07	0.15	0.16	0.88	−0.06	0.07	0.09	0.84

the SE of the estimate obtained in the mixed model. The IPWA estimate also has the smallest root mean square error (RMSE). For large samples, its coverage probabilities (CP) are very close to the nominal level of 95%. These findings are true even for large percentages of censored data (40%) and match the findings in [12]. Whereas the results for a sample size of 500 are not presented, the overall pattern was the same. Again, the estimates from the RT model and the WMM are biased and have larger RMSE.

In our mis-specified model ( $r$ WRT<sub>mis</sub>), a small perturbation of the IPWA has little effect on the estimation and inference. The log-likelihood functions are well approximated by the quadratic functions, and the scores of the parameters from the IPWA and mis-specified models are normally distributed (Shapiro–Wilk tests,  $p > 0.05$ ). As an example, considering 40% censored and 25% missing observations with  $N = 250$ , the scores,  $S(\beta_1)$ , from the IPWA and mis-specified models are distributed as normal with mean zero and Fisher information 67.81 and 67.96, respectively. Because the log-likelihood functions are approximated by their maximum and the curvature at the maximum, it indicates that the likelihood functions are well behaved and have a regular shape.

We also studied the impact of large weights and a mis-specified functional form of the response model on the estimation and inference separately. In this exercise, we found that estimates based on large weights (even for large samples) are biased, and the coverage probabilities are low. For the mis-specified functional form of the response model, the estimates are severely biased, and the coverage probabilities are zero irrespective of the sample sizes that have been considered.

Next, we examine the results of the analysis due to variations in the scenario, the percentage of censoring, and the missing compositions. In Table II, we have presented the simulation results for scenario S3, where 60% of the subjects have missing data because of dropout and 10% have missing data because of death. Again, there are biases associated with all of the parameter estimates with the exception of estimates from the model  $a$ WRT<sub>ipwa</sub> and its mis-specified version. The efficiency of the estimates increased approximately 50% as the sample size increased from 250 to 1000. The RMSE for the parameter estimate associated with the model  $a$ WRT<sub>ipwa</sub> is the smallest among all of the models considered. Simulation results obtained from scenario S2 (not presented), where 30% of the subjects have missing data because of dropout and 10% have missing data because of death, are similar to those obtained from the other two scenarios.



**Table II.** Analysis of simulated non-ignorable missing and left-censored longitudinal data considering 30% completers, 60% dropouts, and 10% subjects died during the study. Only results related to the covariate  $X$  are presented in the table, and the true value of the coefficient is 0.5.

Data feature and model			$N = 250$				$N = 1000$			
Censored	Missing	Model	Bias	SE	RMSE	CP	Bias	SE	RMSE	CP
No	No	iMM	−0.01	0.09	0.09	0.94	0.00	0.04	0.04	0.96
25%	25%	iRT	−0.12	0.09	0.15	0.71	−0.12	0.04	0.13	0.13
		iWRT <sub>1</sub>	−0.09	0.09	0.12	0.81	−0.09	0.04	0.10	0.44
		aWRT <sub>1</sub>	−0.09	0.09	0.12	0.81	−0.09	0.04	0.10	0.44
		aWRT <sub>ipw</sub>	−0.05	0.09	0.10	0.89	−0.08	0.04	0.09	0.62
		aWRT <sub>ipwa</sub>	0.00	0.10	0.10	0.93	0.00	0.05	0.05	0.95
		aWRT <sub>mis</sub>	−0.01	0.10	0.10	0.93	0.00	0.05	0.05	0.95
		rWRT <sub>mis</sub>	−0.01	0.11	0.11	0.91	0.00	0.05	0.05	0.95
		aWRT <sub>large</sub>	0.09	0.11	0.14	0.82	0.10	0.05	0.12	0.50
		aWRT <sub>y-mis</sub>	−0.35	0.03	0.35	0.00	−0.34	0.02	0.34	0.00
		aWMM <sub>h</sub>	−0.07	0.09	0.11	0.85	−0.08	0.04	0.09	0.54
		aWMM <sub>r</sub>	−0.06	0.09	0.11	0.88	−0.07	0.05	0.09	0.65
40%	50%	iRT	−0.08	0.10	0.13	0.88	−0.06	0.05	0.08	0.77
		iWRT <sub>1</sub>	−0.06	0.10	0.12	0.89	−0.05	0.05	0.07	0.83
		aWRT <sub>1</sub>	−0.06	0.10	0.12	0.89	−0.05	0.05	0.07	0.83
		aWRT <sub>ipw</sub>	−0.06	0.10	0.12	0.90	−0.05	0.05	0.07	0.84
		aWRT <sub>ipwa</sub>	0.00	0.11	0.11	0.94	0.00	0.05	0.05	0.95
		aWRT <sub>mis</sub>	0.00	0.11	0.11	0.91	0.00	0.05	0.05	0.93
		rWRT <sub>mis</sub>	0.00	0.12	0.12	0.93	0.00	0.06	0.06	0.95
		aWRT <sub>large</sub>	0.08	0.12	0.14	0.81	0.21	0.07	0.22	0.20
		aWRT <sub>y-mis</sub>	−0.35	0.03	0.35	0.00	−0.34	0.02	0.34	0.00
		aWMM <sub>h</sub>	−0.07	0.10	0.12	0.87	−0.06	0.05	0.08	0.78
		aWMM <sub>r</sub>	−0.03	0.12	0.12	0.94	−0.02	0.06	0.06	0.92

When comparing the estimates across the three simulation scenarios (S1 with 40% dropouts; S2 with 30% dropouts and 10% deaths; S3 with 60% dropouts and 10% deaths), we found that the WRT model based on the IPWA provides unbiased estimates. We also found that for a given sample size, the RMSE and coverage probability of the parameter estimates are very similar across the scenarios. The models that accounted for the censoring through use of tobit regression performed better than models where censored observations were replaced by a randomly imputed value or by half of the lower limit of detection. Overall, the weighted random effects tobit model with IPWA weights produced the best estimates, even with large percentages of missing and censored observations, making this method of choice for analyzing non-ignorable missing and left-censored longitudinal data.

#### 4. Application: analysis of left-censored and non-ignorable missing longitudinal IL6 biomarker data

##### 4.1. Genetic and inflammatory markers of sepsis data, models and results

The GenIMS study was a longitudinal cohort study that was designed to gain an understanding into the role of both genetic and inflammatory biomarkers in the development of sepsis [24]. Because sepsis can result from multiple illnesses, the study focused on recruiting patients with community-acquired pneumonia in order to insure a relatively homogenous group with respect to susceptibility to sepsis. A total of 2320 patients were enrolled into the study through the emergency departments in 28 hospitals in Pennsylvania, Connecticut, Michigan, and Tennessee (2001–2003). The biomarkers were collected daily throughout the first 7 days of hospitalization and may be missing because of death in the hospital during the first 7 days, discharge from the hospital before day 7, or for administrative reasons. In these settings, it is likely that the pattern of any given biomarker will differ depending on the reason for the missing data. Note that there are some intermittent missing data because of administrative reasons, and we are assuming that these intermittently missing data are missing completely at random.

One of the markers of greatest interest in the GenIMS was the pro-inflammatory marker interleukin-6 (IL6); however, the measurement of IL6 was limited because of the sensitivity of the assay resulting in left-censoring of the measure at the lower limit of detection. Figure 1 presents plots of IL6 over the 7 days of measurement for three groups of subjects; those with complete observations over days 1 to 7 (completers), those who have incomplete observations because of hospital discharge (dropout), and those who have incomplete observations because of death during the first 7 days of hospitalization (death). This plot points to the differences in these groups and the need to account for this in the analysis as illustrated by the fact that subjects who died had the highest IL6 levels during the study. In this analysis, we have 598 subjects with complete IL6 measurements, 1119 subjects with incomplete IL6 measurements because of hospital discharge, and 32 subjects with incomplete data because of death during the first 7 days of hospitalization.

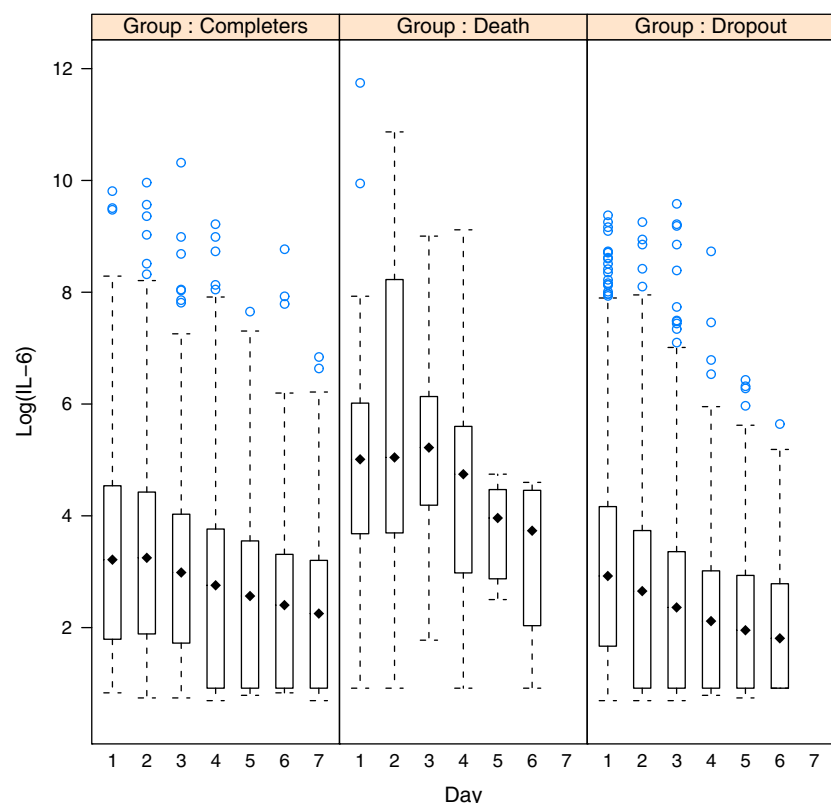
For the comparison of several models that we described in Section 2, we have considered the following random intercept and random slope linear mixed model:

$$IL6_{ij} = \alpha + \beta_0 \text{Day}_{ij} + \beta_1 \text{Age}_i + \beta_2 \text{Statin}_i + \beta_3 \text{APACHE}_i + b_{0i} + b_{1i} \text{Day}_{ij} + \varepsilon_{ij}. \quad (17)$$

The covariates considered in this analysis are day (time) in the study, age of participants, inpatient statin use (binary), and APACHE (acute physiologic and chronic health evaluation). Following Hogan *et al.* [25], we have re-scaled the covariates day ( $\text{day} \div \text{range}(\text{day})$ ), age ( $\text{age} \div 10$  years), and APACHE ( $\text{APACHE} \div \text{range}(\text{APACHE})$ ). After rescaling these covariates, the parameter estimates become more stable by increasing the SE away from zero. This rescaling also improves the interpretation of the estimates. The random effects  $(b_{0i}, b_{1i})^T \sim N(0, D)$ , where  $D$  is a  $2 \times 2$  unstructured covariance matrix and  $\varepsilon_{ij}$  are i.i.d.  $N(0, \sigma^2)$ . The probability of an IL6 value being missing is computed from a logistic regression model that includes the following covariates: logarithm of IL6, age, statin use, pneumonia severity index (PSI), and APACHE, which is a measure of illness severity:

$$\text{logit}(\pi_{ij}) = \lambda_0 + \lambda_{11} \text{Age}_i + \lambda_{12} \text{Statin}_i + \lambda_{13} \text{PSI}_i + \lambda_{14} \text{APACHE}_i + \lambda_2^T \overline{IL6}_{ij}^o + \lambda_3 IL6_{ij}^m \quad (18)$$

where  $\overline{IL6}_{ij}^o = \log(IL6_{i1}^o, IL6_{i2}^o, \dots, IL6_{i,j-1}^o)$ , and  $IL6_{ij}^m$  represents the unobserved IL6 at  $j$ th day. The weights are then computed by inverting the probabilities estimated from this model. For this example,



**Figure 1.** Comparison of IL6 levels by the follow-up category.

we fitted all of the models that have been described in Section 2, with the exception of the complete data model  $i$ MM. Again, we have computed the SE of the estimates by using either the Fisher information matrix, the sandwich estimator, or the asymptotic PL theory.

Table III presents the results obtained from the analysis of the non-ignorable missing and left-censored longitudinal IL6 biomarker data. The  $z$ -values for the intercept ( $> 6.53$ ) and covariate day ( $> -17.42$ ) are large, and because of the space limitations, we have omitted presenting these results. The estimates obtained from the RT models that ignore the non-ignorable missingness are different from the estimates based on the weighting methodology. For example, age and statin use are stronger predictors ( $p$ -value  $\leq 0.02$ ) of IL6 when the estimation is based on the WRT model that uses IPWA weights. These predictors are either marginally significant ( $p$ -value = 0.07) or not significant ( $p$ -value = 0.22) when the estimation is based on the RT model. The estimates of the coefficients and their SEs using WRT models are generally larger than the estimates and SEs obtained from the RT model. This highlights how different models for non-ignorable missing and left-censored data can produce discernibly different, and possibly conflicting, estimates of the effects. Note that for this large data set the SE of the estimate computed by the PL theory is approximately equal to the SE of the estimate computed by the information matrix.

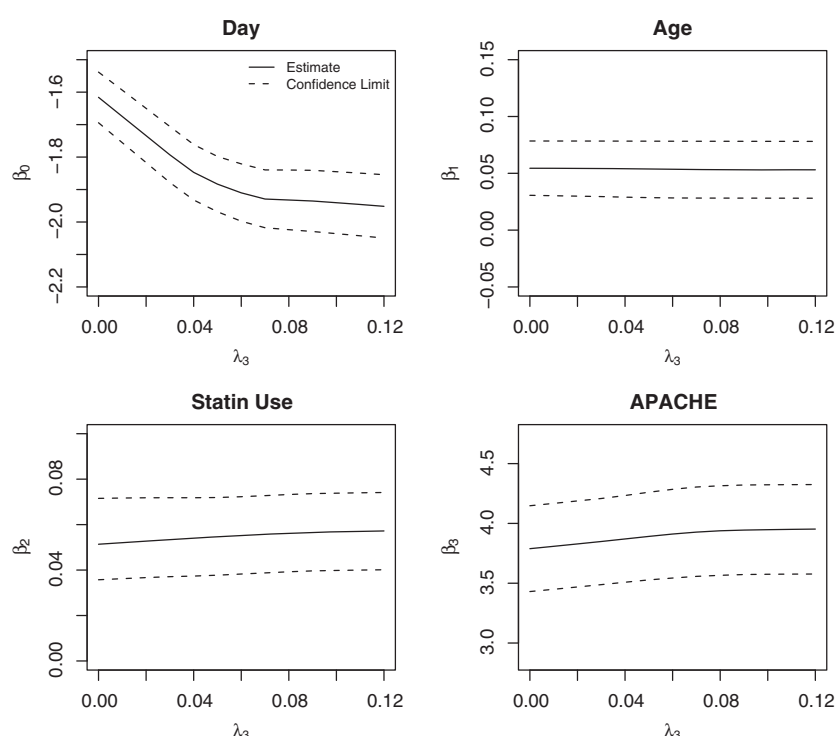
#### 4.2. Sensitivity analysis

Analysis of non-ignorable missing GenIMS longitudinal data necessarily depends on unverifiable assumptions that cannot be tested using the data under analysis, and hence it is necessary to perform

Covariate	Model	Estimates	SE	$z$ -value	$p$ -value
Age	$i$ RT	0.028	0.019	1.48	0.07
	$i$ WRT <sub>1</sub>	0.029	0.021	1.40	0.08
	$a$ WRT <sub>1</sub>	0.029	0.021	1.41	0.08
	$a$ WRT <sub>ipw</sub>	0.054	0.024	2.22	0.01
	$a$ WRT <sub>ipwa</sub>	0.053	0.025	2.12	0.02
	$r$ WRT <sub>mis</sub>	0.052	0.025	2.09	0.02
	$a$ WRT <sub>mis</sub>	0.052	0.026	2.05	0.02
	$a$ WRT <sub>large</sub>	0.051	0.025	2.00	0.02
	$a$ WRT <sub>y-mis</sub>	0.054	0.025	2.15	0.02
	$a$ WMM <sub>h</sub>	0.055	0.020	2.72	0.00
	$a$ WMM <sub>r</sub>	0.058	0.019	3.03	0.00
	$i$ RT	0.008	0.010	0.77	0.22
	$i$ WRT <sub>1</sub>	0.051	0.014	3.56	0.00
Statin use	$a$ WRT <sub>1</sub>	0.051	0.014	3.57	0.00
	$a$ WRT <sub>ipw</sub>	0.051	0.016	3.20	0.00
	$a$ WRT <sub>ipwa</sub>	0.057	0.017	3.42	0.00
	$r$ WRT <sub>mis</sub>	0.058	0.017	3.45	0.00
	$a$ WRT <sub>mis</sub>	0.058	0.017	3.35	0.00
	$a$ WRT <sub>large</sub>	0.060	0.017	3.50	0.00
	$a$ WRT <sub>y-mis</sub>	0.058	0.017	3.44	0.00
	$a$ WMM <sub>h</sub>	0.032	0.013	2.47	0.01
	$a$ WMM <sub>r</sub>	0.029	0.013	2.30	0.01
	$i$ RT	3.028	0.262	11.55	0.00
	$i$ WRT <sub>1</sub>	3.348	0.299	11.21	0.00
	$a$ WRT <sub>1</sub>	3.348	0.298	11.24	0.00
	$a$ WRT <sub>ipw</sub>	3.790	0.359	10.55	0.00
APACHE	$a$ WRT <sub>ipwa</sub>	3.951	0.374	10.56	0.00
	$r$ WRT <sub>mis</sub>	3.966	0.376	10.55	0.00
	$a$ WRT <sub>mis</sub>	3.966	0.422	9.40	0.00
	$a$ WRT <sub>large</sub>	4.021	0.381	10.544	0.00
	$a$ WRT <sub>y-mis</sub>	6.451	0.723	8.918	0.00
	$a$ WMM <sub>h</sub>	3.213	0.300	10.72	0.00
	$a$ WMM <sub>r</sub>	3.132	0.286	10.95	0.00

a sensitivity analysis. On the basis of simulation experience and subject matter expert's judgment in analyzing GenIMS data using the proposed IPWA-based WRT model, we have used an augmented value of 0.12 in the estimation of the continuation probability. This augmentation term is considered as fixed and known, and consequently the parameters in the missing data model and the response model are identified in the sense that the parameters in these models are uniquely determined by the choice of an augmented value (theorem 3, p.1332) [19]. If, for a given choice of augmented value, the parameters in the missing data model (18) are identified and hence the response model parameter estimates are approximately consistent, then the inference is stable. Thus, a sensible strategy is to perform a sensitivity analysis on the choice of the augmentation term,  $\lambda_3$ . We will consider a series of plausible values, including zero, for the augmentation term in the missing data model to see the variation of the parameter estimates in the response model. Again, these augmented values are considered as fixed and known, and then estimation of all other  $\lambda$ 's in the missing data model (18) is based on IPWA estimating Equation (11). If the parameter estimates in model (17) do not change substantially for a range of augmented values, then interpretation of the results would be straightforward. Otherwise, estimation and inference based on the proposed model is not. Note that an augmented value of zero leads to IPW estimation, and hence the missing data mechanism is considered ignorable. In the local influence diagnostic method, this model is called a null model and is used for assessing the influence of unobserved measurements in the analysis by a small perturbation around the null model. Varying the range of  $\lambda_3$  locally around the value for ignorable missingness allows us to study whether small departures from ignorability result in large deviations. This is, broadly, related to local influence type approaches, as proposed by Verbeke et al [26]. Others, such as Copas and Eguchi [27] and Troxel and co-workers [28], have also contributed to this body of research. Although beyond the scope of the current paper, such ideas can be used to develop formal sensitivity tools.

Results of the sensitivity analysis are presented in Figure 2: solid and dash lines indicate the point estimate and confidence limit( $\beta \pm \text{SE}(\hat{\beta})$ ), respectively. Estimates of the fixed effects parameters related to time, age, statin use, and APACHE are presented for varying levels of augmented values, including zero. Because the elevated level of inflammation indicates sepsis and consequently staying in the hospital until resolving the inflammation or until death, a plausible association for the likelihood of observing a biomarker is positively associated with higher levels of IL6 which implies  $\lambda_3 \geq 0$ . Now, by considering  $\lambda_3 \in \mathbb{R}^+$ , the estimate of  $\beta_0$  (panel Day) changes with changing values of  $\lambda_3 < 0.07$ , but, as is shown in



**Figure 2.** Sensitivity of the regression coefficients for various augmented values.

the graph for  $\lambda_3 \geq 0.07$ , the estimate behaves stably. Although there is some variation in the estimates of  $\beta_2$  and  $\beta_3$  for smaller augmented values, these estimates stabilize for  $\lambda_3 \geq 0.07$ . Note that there is no change in the estimation of  $\beta_1$  because of the perturbation of the augmentation term. This supports the argument that the objective function which was used in estimating the parameters and making further inference was robust to small deviations in the selection bias or augmented values.

## 5. Discussion

We proposed an extension of current models to obtain the WRT model for analyzing non-ignorable missing and left-censored longitudinal data. In this study, we have incorporated the IPWA methodology into the likelihood-based inference via weighted pseudo-likelihood theory by taking advantage of the IPWA estimating function that provided a smooth parameter estimate (estimable at a rate of  $\sqrt{N}$ ) for the missing data mechanism [2]. In the weighted pseudo-likelihood-based estimation and inference, weights are considered as nuisance parameters and replaced by consistent estimates in the score function of the parameter of interest. The SEs of the parameter estimates are also corrected by pseudo-likelihood theory. By using the proposed WRT model, we identify the parameters and avoid the computational challenges that arise in the full likelihood-based inference of non-ignorable missing longitudinal data. Using simulation studies, we have shown that the random effect tobit model (assume MAR) and weighted mixed models (NMAR model) that are fitted by replacing the censored values with a half of the detection limit or randomly imputed values produce biased estimates in the presence of non-ignorable missing data. Our proposed WRT model based on the IPWA methodology produced consistently unbiased estimates, and its coverage probability also has the nominal level of 95%. This is true for the different scenarios and sample sizes and for different percentages of missing and censored data. In spite of these attractive features of the proposed WRT model, it should be noted that the proposed model for analyzing non-ignorable missing and left-censored longitudinal data does not resolve the issues of a mis-specified functional form of the outcome variable. The user needs to be very well aware of this, and future investigation should address such sensitivities.

We applied our methodology to the biomarker data obtained from the GenIMS study, a longitudinal cohort study. Analysis of the IL6 longitudinal biomarker data from the GenIMS study shows that the estimates of the coefficients and their SEs vary across the models. Depending on the model used, the inference can be substantially different and possibly conflicting for a covariate that has been used in the model. From the sensitivity analysis of the fitted model, we have seen that the estimates and inferences related to the parameters of interest are robust to the plausible range of augmented values of the missing data model. Finally, from our development and application, we can conclude that the proposed WRT model enjoys the features of the likelihood-based inference with computational simplicity. The model can also be implemented easily for the case of right or doubly censored longitudinal continuous data subject to non-ignorable missingness using standard statistical software.

## Acknowledgements

We thank Dr Derek Angus and the CRISMA laboratory for access to the GenIMS data. We are indebted to the nurses, respiratory therapists, phlebotomists, physicians, and other health-care professionals, as well as the patients and their families who supported this trial. A complete list of GenIMS investigators is available at [www.ccm.upmc.edu/genims](http://www.ccm.upmc.edu/genims) investigators. The GenIMS study was funded via grant R01 GM61992 by the National Institute of General Medical Sciences.

## References

1. Kenward MG, Carpenter JR. Multiple imputation. In *Longitudinal Data Analysis*, Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G (eds). GCRG: London, 2009.
2. Rotnitzky A. Inverse probability weighted methods. In *Longitudinal Data Analysis*, Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G (eds). CRC: London, 2009.
3. Philipson PM, Ho WK, Henderson R. Comparative review of methods for handling drop-out in longitudinal studies. *Statistics in Medicine* 2008; **27**:6276–6298.
4. Little RJ. Selection and pattern-mixture models. In *Longitudinal Data Analysis*, Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G (eds). CRC Press: New York, 2009.
5. Albert PS, Follmann DA. Shared-parameter models. In *Longitudinal Data Analysis*, Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G (eds). CRC: London, 2009.



6. Yuan Y, Little RJA. Mixed-effect hybrid models for longitudinal data with nonignorable dropout. *Biometrics* 2009; **65**:478–486.
7. Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G. *Longitudinal Data Analysis*. Chapman & Hall/CRC: Boca Raton, FL, 2009.
8. Horvitz DG, Thompson DJ. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 1952; **47**:663–685.
9. Robins JM, Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology—Methodological Issues*, Jewell N, Dietz K, Farewell V (eds). Birkhäuser: Boston, MA, 1992.
10. Wu L. *Mixed Effects Models for Complex Data*. CRC Press: New York, 2010.
11. Tobin J. Estimation of relationship for limited dependent variables. *Econometrica* 1958; **26**:24–36.
12. Epstein MP, Lin X, Boehnke M. A tobit variance-component method for linkage analysis of censored trait data. *American Journal of Human Genetics* 2003; **72**:611–620.
13. Huges J. Mixed effects models with censored data with applications to HIV RNA levels. *Biometrics* 1999; **55**:625–629.
14. Jacqmin-Gadda H, Thiebaut R, Chene G, Commenges D. Analysis of left-censored longitudinal data with application to viral load in HIV infection. *Biostatistics* 2000; **1**:355–368.
15. Lyles RH, Lyles CM, Taylor DJ. Random regression models for human immunodeficiency virus ribonucleic acid data subject to left censoring and informative drop-outs. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 2000; **49**:485–497.
16. Gao S, Thiebaut R. Mixed effect models for truncated longitudinal outcomes with nonignorable missing data. *Journal of Data Science* 2009; **7**:13–25.
17. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982; **38**:963–974.
18. Gong G, Samaniego FJ. Pseudo maximum likelihood estimation: theory and applications. *The Annals of Statistics* 1981; **9**:861–869.
19. Rotnitzky A, Robins JM, Scharfstein DO. Semiparametric regression for repeated outcomes with nonignorable nonresponse. *Journal of the American Statistical Association* 1998; **93**:1321–1339.
20. Lawless JF, Kalbfleisch JD, Wild CJ. Semiparametric methods for response-selective and missing data problems in regression. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 1999; **61**:413–438.
21. Kang JDY, Schafer JL. Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science* 2007; **22**:523–539.
22. Robins J, Sued M, Lei-Gomez Q, Rotnitzky A. Comment: performance of double-robust estimators when “inverse probability” weights are highly variable. *Statistical Science* 2007; **22**:544–559.
23. Demirtas H, Hedeker D. Gaussianization-based quasi-imputation and expansion strategies for incomplete correlated binary responses. *Statistics in Medicine* 2007; **26**:782–799.
24. Kellum JA, Kong L, Fink MP, Weissfeld LA, Yealy DM, Pinsky MR, Fine J, Krichevsky A, Delude RL, Angus DC. Understanding the inflammatory cytokine response in pneumonia and sepsis: results of the Genetic and Inflammatory Markers of Sepsis (GenIMS) Study. *Archives of Internal Medicine* 2007; **167**:1655–1663.
25. Hogan JW, Lin X, Herman B. Mixtures of varying coefficient models for longitudinal data with discrete or continuous nonignorable dropout. *Biometrics* 2004; **60**:854–864.
26. Verbeke G, Molenberghs G, Thijs H, Lesaffre E, Kenward MG. Sensitivity analysis for nonrandom dropout: a local influence approach. *Biometrics* 2001; **57**:7–14.
27. Copas J, Eguchi S. Local sensitivity approximations for selectivity bias. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2001; **63**:871–895.
28. Troxel AB, Harrington DP, Lipsitz SR. Analysis of longitudinal data with non-ignorable non-monotone missing values. *Journal of the Royal Statistical Society Series C-Applied Statistics* 1998; **47**:425–438.