Title

Estimation and inference on correlations between biomarkers with repeated measures and left-censoring due to minimum detection levels

Permalink

https://escholarship.org/uc/item/33g5s1qf

Journal

Statistics in Medicine, 31(21)

ISSN

0277-6715

Authors

Xie, Xianhong
Xue, Xiaonan
Gange, Stephen J
et al.

Publication Date

2012-09-20

DOI

10.1002/sim.5371

Peer reviewed

Estimation and inference on correlations between biomarkers with repeated measures and left-censoring due to minimum detection levels

Xianhong Xie[a,*,†], **Xiaonan Xue**[a], **Stephen J. Gange**[b], **Howard D. Strickler**[a], **Mimi Y. Kim**[a], and **WIHS HPV Study Group**

[a]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, U.S.A

[b]Department of Epidemiology, John Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, U.S.A

Abstract

Statistical approaches for estimating and drawing inference on the correlation between two biomarkers which are repeatedly assessed over time and subject to left-censoring due to minimum detection levels are lacking. We propose a linear mixed-effects model and estimate the parameters with the Monte Carlo Expectation Maximization (MCEM) method. Inferences regarding the model parameters and the correlation between the biomarkers are performed by applying Louis's method and the delta method. Simulation studies were conducted to compare the proposed MCEM method with existing methods including the MLE method, the multiple imputation (MI) method, and two widely used ad hoc approaches: replacing the censored values with the detection limit (DL) or with half of the detection limit (HDL). The results show that the performance of the MCEM with respect to relative bias and coverage probability for the 95% confidence interval is superior to the DL and HDL approaches and exceeds that of the MI method at medium to high levels of censoring, and the standard error estimates from the MCEM method are close to ideal. The MLE method can estimate the parameters accurately; however, a non-positive definite information matrix can occur so that the variances are not estimable. These five methods are illustrated with data from a longitudinal HIV study to estimate and draw inference on the correlation between HIV RNA levels measured in plasma and in cervical secretions at multiple time points.

Keywords

information matrix; longitudinal data; mixed-effects; monte carlo expectation maximization

1. Introduction

Biomarkers of disease are often measured repeatedly over time in epidemiological studies. In addition, the measurements may be left-censored due to minimum detection levels below which the level of the biomarker cannot be quantified. These issues can complicate estimation of the correlation between two different biomarkers measured in the same subject or between levels of the same biomarker measured in two different biological specimens

[*]Correspondence to: Xianhong Xie, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Ave, Belfer Bldg. #1308, Bronx, NY 10461, U.S.A.
[†]xianhong.xie@einstein.yu.edu

(e.g., tissue and blood). Even in the absence of left-censoring, the simple Pearson correlation coefficient cannot be applied to such data because repeated measures from the same subject tend to be correlated [1]. Hamlett *et al.* [2] proposed the use of a mixed-effects model to estimate the correlations involving repeated measures, but these approaches cannot accommodate left-censored observations.

In the example which motivated the methods in this paper, HIV RNA levels were assessed in both cervical secretion and plasma specimens and levels below the limit of detection were not measurable [3, 4]. Due to biologic compartmentalization, the two values are expected to differ [5], but the level of correlation is a matter of research interest that is obscured by problems caused by the lower limit of detectability. To address left-censored data of this type, two ad hoc approaches have been widely applied: replacing the censored values with the detection limit (DL) or with half of the detection limit (HDL), or with some other arbitrary value related to the DL. However, the validity of these ad hoc approaches is questionable since the variability of the observations is artificially reduced. Another potential approach for handling left-censored data is multiple imputation (MI) [6, 7]. Rather than filling in the censored observations with a single value, an imputation model is used to impute multiple values for the censored observations and the usual analysis is then performed with the multiply imputed "complete" data sets.

With a single biomarker subject to left-censoring, a maximum likelihood (ML) approach has been proposed in which the censoring information is incorporated into the likelihood of the observed data [8–10] to estimate the association of the biomarker (i.e., HIV RNA level) with exposure to treatment (i.e., treatment with anti-retroviral therapy). However, little research has been conducted to address associations/correlations between bivariate repeated measurements (e.g., HIV RNA level in two types of specimens) with censoring. Hamlett *et al.* [2] studied the multivariate repeated measurements problem without censoring by maximizing the likelihood, but a method for inferences regarding the correlation parameters was not addressed. Roy [11] further generalized the method in [2] to allow the measurement errors to be correlated over time. Lyles *et al.* [12] studied the bivariate censoring problem without repeated measurements by optimizing the likelihood and drawing inferences with a profile likelihood approach. Hughes [8] studied the repeated measures with censoring problem for a single biomarker but the inferences focused only on the fixed effect parameters. Thiebaut *et al.* [13] studied bivariate correlation between CD4 count and HIV RNA viral load with informative drop-out among HIV sero-positive patients, however only one biomarker (i.e. HIV RNA viral load) was subject to left-censoring and the focus was on estimating HIV treatment effect, also it was assumed that correlations between biomarkers were completely explained by random effects for each biomarker and there was no correlation between measurement errors. Albert [14] studied the modeling of two repeatedly measured simian immunodeficiency virus (SIV) viral loads using two different assays with a single random intercept for both viral loads, and the focus was on finding the best design strategy to selectively perform the more accurate and expensive SIV viral load assay only on a subset of individuals so as to reduce cost and maintain accuracy. When the availability of two assays is not an issue (e.g. pilot study) and the focus is on estimating and inference on correlations, then more accurately modeling the covariance structure becomes important; it is desirable to allow the covariance structure to be flexible [2, 11] and use the data to estimate the parameters in covariance under the general framework.

To our knowledge, only Hopke *et al.* [15] addressed multivariate repeated measures with left censoring and relatively general covariance structure and proposed a multiple imputation algorithm. However, left-censoring due to a minimum detection level does not satisfy the "missing at random" condition [16] typically assumed in multiple imputation since an observation is left-censored only when the value is lower than the detection limit; hence the

pattern of missingness depends on the missing values, and the performance of the multiple imputation in the case of left-censoring needs to be examined. In this paper, we propose a maximum likelihood approach to address the bivariate repeated measures censoring problem using Monte-Carlo Expectation Maximization (MCEM) and compare it with the two common ad hoc approaches, DL and HDL, and the MI approach [15]. An alternative to the MCEM method is the MLE method in which the likelihood function is approximated by a quadrature and then optimized using a modified Newton-Raphson algorithm [9]. The observed information matrix is obtained with numerical differentiation. We also compare the MLE method to our MCEM approach.

We first briefly review the multivariate mixed-effects model, then describe estimation and inference procedures for uncensored data, as well as for censored data based on the DL, HDL, MI, MLE, and the proposed MCEM methods. Details of the derivations of the formulas in the methods section are given in the electronic appendix. The performances of these five methods: DL, HDL, MI, MLE, and MCEM, are compared via simulation studies. The methods are applied to data from a longitudinal HIV RNA study which involved the estimation and inference of correlations between HIV RNA levels in cervical secretions and in plasma measured repeatedly over time. Finally, we discuss limitations and possible extensions of the MCEM approach.

2. Methods

2.1. Multivariate mixed-effects model

Consider the following multivariate mixed-effects model [17]:

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \boldsymbol{Z}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i, \quad (1)$$

where $\mathbf{y}_i$ is a $n_i \times 2$ matrix (in the bivariate case) for the ith subject ($i = 1, \cdots, m$), n_i is the number of repeated measurements (or time points) for subject i, $\mathbf{X}_i$ and $\mathbf{Z}_i$ are both $n_i \times 1$ matrices consisting of all 1's. The fixed effects $\boldsymbol{\beta} = (\beta_1, \beta_2)$ and the random effects $\mathbf{b}_i = (b_{i1}, b_{i2})$ are row vectors of length two, and the measurement error $\boldsymbol{\varepsilon}_i$ is a $n_i \times 2$ matrix, $\boldsymbol{\varepsilon}_i = (\boldsymbol{\varepsilon}_{i1}^T, \cdots, \boldsymbol{\varepsilon}_{in_i}^T)^T$, T is the transpose operator. We also assume $\mathbf{b}_i$ and $\boldsymbol{\varepsilon}_i$ are normally distributed, independent and the rows of $\boldsymbol{\varepsilon}_i$ are mutually independent so that

$$\text{vec}(\mathbf{b}_i) = \begin{pmatrix} b_{i1} \\ b_{i2} \end{pmatrix} \sim N(\mathbf{0}, \boldsymbol{\Psi}), \quad (2)$$
$$\text{vec}(\boldsymbol{\varepsilon}_i) \sim N(\mathbf{0}, \textstyle\sum \otimes \mathbf{I}_{n_i}),$$

where $\text{vec}(\cdot)$ is the vectorization operator which stacks matrix by columns:

$\text{vec}((\mathbf{a}_1 \cdots \mathbf{a}_n)) = (\mathbf{a}_1^T \cdots \mathbf{a}_n^T)^T$, $\mathbf{a}_i$, $1 \leq i \leq n$ are column vectors of the same length, $\boldsymbol{\Psi}$ and Σ are both 2×2 positive definite matrices, $\otimes$ is the kronecker product, and $\mathbf{I}_n$ is the identity matrix of dimension $n \times n$. We have

$$\text{vec}(\mathbf{y}_i) \sim \text{N}(\tilde{\mathbf{X}}_i \, \text{vec}(\boldsymbol{\beta}), \mathbf{V}_i), \quad (3)$$

where $\tilde{\mathbf{X}_i} = \mathbf{I}_2 \otimes \mathbf{X}_i$, $\mathbf{V}_i = \tilde{\boldsymbol{Z}}_i \boldsymbol{\Psi} \tilde{\boldsymbol{Z}}_i^T + \sum \otimes \mathbf{I}_{n_i}$, $\tilde{\mathbf{Z}_i} = \mathbf{I}_2 \otimes \mathbf{Z}_i$. Let

$$\mathbf{\Psi} = \begin{pmatrix} \sigma_{r11}^2 & \sigma_{r12} \\ \sigma_{r12} & \sigma_{r22}^2 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} \sigma_{e11}^2 & \sigma_{e12} \\ \sigma_{e12} & \sigma_{e22}^2 \end{pmatrix},$$

and denote the (j, k)-th element of $\mathbf{y}_i$ by y_{ijk} for $j=1, \cdots, n_i$; $k=1, 2$. The above mixed-effects model implies the following covariance structure for $\mathbf{y}_i$,

$$\begin{aligned}
\mathrm{Cov}(y_{ij1}, y_{ij1}) &= \sigma_{r11}^2 + \sigma_{e11}^2, \\
\mathrm{Cov}(y_{ij2}, y_{ij2}) &= \sigma_{r22}^2 + \sigma_{e22}^2, \\
\mathrm{Cov}(y_{ij1}, y_{ij2}) &= \sigma_{r12} + \sigma_{e12}, \\
\mathrm{Cov}(y_{ij1}, y_{ij'1}) &= \sigma_{r11}^2, \\
\mathrm{Cov}(y_{ij2}, y_{ij'2}) &= \sigma_{r22}^2, \\
\mathrm{Cov}(y_{ij1}, y_{ij'2}) &= \sigma_{r12}, \\
1 \le j &\ne j' \le n_i.
\end{aligned}$$

The model in Hughes [8] can only handle the case when the measurement error variances σ_{e11}^2 and σ_{e22}^2 are the same and the measurement error covariance σ_{e12} is 0; the approach in Thiebaut *et al*. [13] restricts the measurement error covariance σ_{e12} to be 0; the approach in Albert [14] restricts the random effect variances σ_{r11}^2 and σ_{r22}^2 to be the same and the random effect covariance σ_{r12} to be 0; with our approach, there is no restrictions on the parameters.

Denote the observed data by $\{(\mathbf{Q}_i, \mathbf{C}_i), 1 \le i \le m\}$,

$$\begin{aligned}
y_{ijk} &< Q_{ijk} \quad \text{if } C_{ijk} = -1, \\
y_{ijk} &= Q_{ijk} \quad \text{if } C_{ijk} = 0, \\
y_{ijk} &> Q_{ijk} \quad \text{if } C_{ijk} = 1, \\
i = 1, \cdots, m, \quad & j = 1, \cdots, n_i, \ k = 1, 2.
\end{aligned}$$

Let $\boldsymbol{\theta} = (\beta_1, \beta_2, \sigma_{r11}^2, \sigma_{r12}, \sigma_{r22}^2, \sigma_{e11}^2, \sigma_{e12}, \sigma_{e22}^2)^T$ and $\tilde{\boldsymbol{\theta}} = (\beta_1, \beta_2, \sigma_{r11}^2, \rho_r, \sigma_{r22}^2, \sigma_{e11}^2, \rho_e, \sigma_{e22}^2)^T$, where $\tilde{\boldsymbol{\theta}}$ differs from $\boldsymbol{\theta}$ only with respect to the parameterization of the correlation parameters: $\rho_r = \sigma_{r12} / \sqrt{\sigma_{r11}^2 \sigma_{r22}^2}$ and $\rho_e = \sigma_{e12} / \sqrt{\sigma_{e11}^2 \sigma_{e22}^2}$. Our goal is to estimate and draw inferences about the parameter vector $\boldsymbol{\theta}$ and the correlation coefficient between the two variables

$$\rho = \mathrm{Corr}(y_{ij1}, y_{ij2}) = \frac{\sigma_{r12} + \sigma_{e12}}{\sqrt{(\sigma_{r11}^2 + \sigma_{e11}^2)(\sigma_{r22}^2 + \sigma_{e22}^2)}}. \quad (4)$$

We assumed a two-level correlation structure between the two biomarkers: subject level and time level. The quantity ρ_r measures the correlation between biomarkers due to subject-specific random variation, ρ_e measures the correlation between biomarkers within the same subject at the same time point due to time-specific random variation, and ρ is the overall

correlation when both subject-specific and time-specific random variations are taken into account.

2.2. Estimation and inference with uncensored data

When the data are uncensored, the log-likelihood has the form

$$ll(\boldsymbol{\theta}) = \sum_{i=1}^{m} \left(-\frac{n_i}{2} \log(2\pi) - \frac{1}{2} \log|\mathbf{V}_i| - \frac{1}{2} \mathbf{u}_i^T \mathbf{V}_i^{-1} \mathbf{u}_i \right), \quad (5)$$

where $\mathbf{u}_i = \text{vec}(\mathbf{y}_i) - \tilde{\mathbf{X}}_i \text{vec}(\boldsymbol{\beta})$. The log-likelihood can be maximized with respect to $\boldsymbol{\theta}$ using the Newton-Raphson algorithm or other numerical optimization algorithm [18]. The observed information matrix $\mathbf{I}_y(\widehat{\boldsymbol{\theta}_{uc}})$ at the maximum likelihood estimate (MLE), $\widehat{\boldsymbol{\theta}_{uc}}$, can be obtained using numeric differentiation for the negative second order derivative of the log-likelihood. For the model considered in this paper, an analytic formula is also available (Equation (A.3) in electronic appendix). Inferences about $\tilde{\boldsymbol{\theta}}$ can be made using the transformed parameters $\boldsymbol{\eta} = g(\boldsymbol{\theta})$, log square root transformation on the variances, and Fisher's z-transformation (without the constant 1/2) for the correlation parameters [18, 19]:

$$\boldsymbol{\eta} = g(\boldsymbol{\theta}) = \left(\beta_1 \quad \beta_2 \quad \log(\sigma_{r11}) \quad \log(\frac{1+\rho_r}{1-\rho_r}) \quad \log(\sigma_{r22}) \quad \log(\sigma_{e11}) \quad \log(\frac{1+\rho_e}{1-\rho_e}) \quad \log(\sigma_{e22}) \right)^T. \quad (6)$$

The advantage of drawing inferences using $\boldsymbol{\eta}$ is that the ranges of the components of $\boldsymbol{\eta}$ are not constrained. The observed information for $\boldsymbol{\eta}$ with the MLE estimate $\widehat{\boldsymbol{\theta}_{uc}}$ is

$$\tilde{\mathbf{I}}_y(\widehat{\boldsymbol{\eta}}_{uc}) = \mathbf{J}(\widehat{\boldsymbol{\eta}}_{uc})^T \mathbf{I}_y(\widehat{\boldsymbol{\theta}}_{uc}) \mathbf{J}(\widehat{\boldsymbol{\eta}}_{uc}) - \left(\mathbf{I}_8 \otimes \mathbf{S}_y(\widehat{\boldsymbol{\theta}}_{uc})^T \right) \frac{\partial \mathbf{J}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \Big|_{\boldsymbol{\eta} = \widehat{\boldsymbol{\eta}}_{uc}}, \quad (7)$$

where $\widehat{\boldsymbol{\eta}}_{uc} = g(\widehat{\boldsymbol{\theta}}_{uc})$, and $\mathbf{J}(\cdot)$ is the Jacobian matrix for the transformation $\boldsymbol{\theta} = g^{-1}(\boldsymbol{\eta})$. $\mathbf{S}_y(\cdot)$ is the gradient of the uncensored log-likelihood (electronic appendix). The variance of $\widehat{\boldsymbol{\eta}}_{uc}$ can be approximated by $\tilde{\mathbf{I}}_y(\widehat{\boldsymbol{\eta}}_{uc})^{-1}$. The $100(1-\alpha)\%$ confidence interval for $\boldsymbol{\eta}$ is given by $\widehat{\boldsymbol{\eta}}_{uc} \pm z_{\alpha/2} \sqrt{\text{diag}(\tilde{\mathbf{I}}_y(\widehat{\boldsymbol{\eta}}_{uc})^{-1})}$, where $z_{\alpha/2}$ is the upper $100*\alpha/2$-th percentile of standard normal distribution. Transforming back to the original scale, $\tilde{\boldsymbol{\theta}}$, yields the corresponding confidence intervals.

For inferences about ρ, to which the Fisher's z-transformation has been applied, we use the delta method [20] to obtain:

$$\text{Var}\left(\log(\frac{1+\widehat{\rho}_{uc}}{1-\widehat{\rho}_{uc}}) \right) \approx h^2 \nabla_{\widehat{\boldsymbol{\theta}}_{uc}}^T \mathbf{I}_y(\widehat{\boldsymbol{\theta}}_{uc})^{-1} \nabla_{\widehat{\boldsymbol{\theta}}_{uc}}, \quad (8)$$

where $\widehat{\rho_{uc}}$ is generated by Equation (4) with $\boldsymbol{\theta}$ replaced by the MLE $\widehat{\boldsymbol{\theta}_{uc}}$,

$$h = \frac{d\log((1+\rho)/(1-\rho))}{d\rho}\Big|_{\rho=\widehat{\rho}_{uc}} = \frac{2}{1-\widehat{\rho}_{uc}^2},$$

$$\nabla_{\widehat{\boldsymbol{\theta}}_{uc}} = \frac{d\rho}{d\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_{uc}} = \left(0 \quad 0 \quad -cd^3 s_2/2 \quad d \quad -cd^3 s_1/2 \quad -cd^3 s_2/2 \quad d \quad -cd^3 s_1/2 \right)^T \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_{uc}},$$

$$c = \sigma_{r12} + \sigma_{e12}, \quad d = 1/\sqrt{(\sigma_{r11}^2 + \sigma_{e11}^2)(\sigma_{r22}^2 + \sigma_{e22}^2)}, \quad s_1 = \sigma_{r11}^2 + \sigma_{e11}^2, \quad s_2 = \sigma_{r22}^2 + \sigma_{e22}^2.$$

Denote the right hand side of Formula (8) by $\hat{v_{uc}}$, the 100(1−α)% confidence interval for $\log((1+\rho)/(1-\rho))$ (denote by $\tilde{\rho}$) is $\log\left((1+\hat{\rho}_{uc})/(1-\hat{\rho}_{uc})\right) \pm z_{\alpha/2} \sqrt{\hat{v}_{uc}}$. The confidence interval for ρ can be obtained by transforming back to the original scale.

With the DL and HDL methods, the censored data are replaced by the values of the detection limits and half of the detection limits, respectively, and the methods above for uncensored data are applied.

2.3. Multiple imputation

Multiple imputation is an alternative approach for handling missing data due to censoring. Hopke *et al.* [15] applied the MI method to impute missing as well as left-censored data due to detection thresholds. The missing data and the left-censored data were imputed differently: the former was imputed based on a normal distribution given the observed data, parameter vector and imputed values for the left-censored data and the latter was based on a truncated normal distribution given the observed data, parameter vector and imputed values for the missing data. We adopted their method to address left-censored data without the presence of other types of missingness. Specifically, a Gibbs sampler method described below was used. Starting from an initial estimate, $\theta^{(0)}$, and initial imputed values for censored biomarkers, the following three steps are performed iteratively:

1. Draw $\mathbf{b}_i^{(t+1)} \sim P\left(\mathbf{b}_i | \mathbf{y}_{obs}, \mathbf{y}_{cens}^{(t)}, \theta^{(t)}\right)$, $i=1, \cdots, m$,

2. Draw $\theta^{(t+1)} \sim P\left(\theta | \mathbf{y}_{obs}, \mathbf{y}_{cens}^{(t)}, \mathbf{b}^{(t+1)}\right)$,

3. Draw $\mathbf{y}_{i(cens)}^{(t+1)} \sim P\left(\mathbf{y}_{i(cens)} | \mathbf{Q}_i, \mathbf{C}_i, \mathbf{b}_i^{(t+1)}, \theta^{(t+1)}\right)$, $i=1, \cdots, m$,

where t indexes the iteration. A set of M (say 5) imputations for the censored $\mathbf{y}$ values are drawn to yield M complete sets of data and the methods described above for uncensored data are applied to each complete data set. The parameter estimates and variance estimates of the transformed parameters η and $\tilde{\rho}$ from the different data sets are then combined as in [21] to obtain the MI estimate, variance estimate, and degrees of freedom.

The introduction of random effects into the Gibbs sampler simplifies the random sampling. Since the conditional distribution in Step 3 is a product of truncated univariate normal distributions and truncated bivariate normal distributions, the censored values can be drawn easily.

2.4. Maximum likelihood estimation method

A MLE method for left-censored HIV viral loads was proposed in Jacqmin-Gadda *et al.* [9]. By re-arranging the order of data in $vec(\mathbf{y}_i)^T$ according to $vec(\mathbf{C}_i)$ to put all observed data before censored data ($\mathbf{y}_i^{oT}, \mathbf{y}_i^{cT}$) for each subject, the permuted underlying variance matrix can be written as

$$\tilde{\mathbf{V}}_i = \begin{pmatrix} \mathbf{V}_i^o & \mathbf{V}_i^{coT} \\ \mathbf{V}_i^{co} & \mathbf{V}_i^c \end{pmatrix},$$

where $\mathbf{V}_i^o$ and $\mathbf{V}_i^c$ are the variance matrices within the observed and the censored data respectively for subject i, and $\mathbf{V}_i^{co}$ is the covariance matrix between the censored and the observed data for subject i. Let the corresponding partitions for permuted $vec(\mathbf{Q}_i)^T$ and $\tilde{\mathbf{X}}_i^T$

be ($\mathbf{Q}_i^{oT}, \mathbf{Q}_i^{cT}$) and ($\tilde{\mathbf{X}}_i^{oT}, \tilde{\mathbf{X}}_i^{cT}$) respectively. Under the assumption of multivariate normal distribution (3) for $vec(\mathbf{y}_i)$, the likelihood of the data is

$$L(\boldsymbol{\theta})=\prod_{i=1}^{m} f_i^o(\mathbf{Q}_i^o|\boldsymbol{\theta})\Phi_i^{c|o}(\mathbf{Q}_i^c|\boldsymbol{\theta}),$$

where f_i^o is the multivariate normal density with mean $\boldsymbol{\mu}_i^o(=\tilde{\mathbf{X}}_i^o vec(\boldsymbol{\beta}))$ and variance $\mathbf{V}_i^o$, and $\Phi_i^{c|o}$ is the normal distribution function for $\mathbf{y}_i^c$ given $\mathbf{y}_i^o$, which involves multivariate integral of normal density,

$$\Phi_i^{c|o}(\mathbf{Q}_i^c|\boldsymbol{\theta})=\int_{-\infty}^{Q_{i1}^c}\cdots\int_{-\infty}^{Q_{in_i^c}^c}(2\pi)^{-n_i^c/2}|\mathbf{V}_i^{c|o}|^{-1/2}\exp(-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu}_i^{c|o})^T\mathbf{V}_i^{c|o^{-1}}(\mathbf{z}-\boldsymbol{\mu}_i^{c|o}))d\mathbf{z},$$

with

$$\boldsymbol{\mu}_i^{c|o}=\tilde{\mathbf{X}}_i^c vec(\boldsymbol{\beta})+\mathbf{V}_i^{co}\mathbf{V}_i^{o^{-1}}(\mathbf{Q}_i^o-\boldsymbol{\mu}_i^o),$$
$$\mathbf{V}_i^{c|o}=\mathbf{V}_i^c-\mathbf{V}_i^{co}\mathbf{V}_i^{o^{-1}}\mathbf{V}_i^{coT},$$

n_i^c is the number of censored data for subject i, $Q_{ik}^c, 1 \le k \le n_i^c$ are the components of $\mathbf{Q}_i^c$. The Fortran subroutine SADMVN by Genz [22] was used to compute the integrals, which works well when the size of integral (i.e. n_i^c) is less than 10. The Marquardt algorithm [23] was used on the log-likelihood with respect to the following parameters: the fixed effect parameters and the Cholesky factorizations of $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$, to find the optimum. The optimization algorithm is stopped if the sum of squares of changes in parameter estimates, the absolute change in log-likelihood, and the normalized gradients are less than some tolerance values or maximum number of iterations has been reached. After stopping of the optimization, the second-order finite difference method on the negative log-likelihood with respect to the original parameter $\boldsymbol{\theta}$ was used to get the observed information matrix. The variance-covariance matrix was estimated by the inverse of the observed information matrix. The inferences on the correlations were done with delta method. One limitation of the MLE method is that the algorithm can generate a non-positive definite estimated information matrix because the use of the quadrature method for the approximation of the multivariate normal cumulative distribution function has limited precision, and the numerical approximation of the second derivative of the log-likelihood function also has error. In the case of a non-positive definite matrix, variance estimates cannot be obtained easily. Other methods such as bootstrapping [24] may be used although they are computationally intensive.

2.5. MCEM

Our MCEM method is as follows: let the complete data be $\{(\mathbf{Q}_i, \mathbf{C}_i, \mathbf{y}_i, \mathbf{b}_i, \boldsymbol{\varepsilon}_i), 1 \le i \le m\}$, the E-step is

$$E\left(\text{vec}(\mathbf{y}_i)|\mathbf{Q}_i,\mathbf{C}_i\right)=E\left(\mathbf{u}_i|\mathbf{Q}_i,\mathbf{C}_i\right)+\tilde{\mathbf{X}}_i\text{vec}(\boldsymbol{\beta}),$$
$$E\left(\text{vec}(\mathbf{b}_i)\text{vec}(\mathbf{b}_i)^T|\mathbf{Q}_i,\mathbf{C}_i\right)=\boldsymbol{\Psi}\tilde{\mathbf{Z}}_i^T\mathbf{W}_iE\left(\mathbf{u}_i\mathbf{u}_i^T|\mathbf{Q}_i,\mathbf{C}_i\right)\mathbf{W}_i\tilde{\mathbf{Z}}_i\boldsymbol{\Psi}+\boldsymbol{\Psi}-\boldsymbol{\Psi}\tilde{\mathbf{Z}}_i^T\mathbf{W}_i\tilde{\mathbf{Z}}_i\boldsymbol{\Psi},$$
$$E\left(\text{vec}(\boldsymbol{\varepsilon}_{ij})\text{vec}(\boldsymbol{\varepsilon}_{ij})^T|\mathbf{Q}_i,\mathbf{C}_i\right)=\sum\mathbf{W}_{i,j.}E\left(\mathbf{u}_i\mathbf{u}_i^T|\mathbf{Q}_i,\mathbf{C}_i\right)\mathbf{W}_{i,j.}^T\Sigma+\Sigma-\sum\mathbf{W}_{i,jj}\Sigma, \quad (9)$$

where the expectations are evaluated at the value of the current parameter estimate for $\boldsymbol{\theta}$, $\mathbf{W}_i=\mathbf{V}_i^{-1}$, $\mathbf{W}_{i,jj}$ is the 2×2 submatrix of $\mathbf{W}_i$ corresponding to the covariance of $\text{vec}(\mathbf{y}_{ij})=(y_{ij1},y_{ij2})^T$, $\mathbf{W}_{i,j}$ is the $2\times 2n_i$ submatrix of $\mathbf{W}_i$ corresponding to the covariance of $\text{vec}(\mathbf{y}_{ij})$ with $\text{vec}(\mathbf{y}_i)$. The second and third equations in (9) are derived by first computing the conditional expectation of $\text{vec}(\mathbf{b}_i)\text{vec}(\mathbf{b}_i)^T$ and $\text{vec}(\boldsymbol{\varepsilon}_{ij})\text{vec}(\boldsymbol{\varepsilon}_{ij})^T$ on $\mathbf{y}_i$ respectively, then computing the expectations conditional on $(\mathbf{Q}_i,\mathbf{C}_i)$. We calculate the first and second order moments of $\mathbf{u}_i$ conditional on $(\mathbf{Q}_i,\mathbf{C}_i)$ as follows,

$$E\left(\mathbf{u}_i|\mathbf{Q}_i,\mathbf{C}_i\right)\approx\sum_{l=1}^{L}\mathbf{u}_i^l/L,$$
$$E\left(\mathbf{u}_i\mathbf{u}_i^T|\mathbf{Q}_i,\mathbf{C}_i\right)\approx\sum_{l=1}^{L}\mathbf{u}_i^l(\mathbf{u}_i^l)^T/L, \quad (10)$$

where $\mathbf{u}_i^l$ is drawn from the truncated normal distribution $p(\mathbf{u}_i|\mathbf{Q}_i,\mathbf{C}_i)$ using a Gibbs sampler [8] and L is the number of draws.

The M-step is,

$$\widehat{\boldsymbol{\beta}}^T=\left(\sum_{i=1}^{m}\tilde{\mathbf{X}}_i^T\mathbf{W}_i\tilde{\mathbf{X}}_i\right)^{-1}\left(\sum_{i=1}^{m}\tilde{\mathbf{X}}_i^T\mathbf{W}_iE\left(\text{vec}(\mathbf{y}_i)|\mathbf{Q}_i,\mathbf{C}_i\right)\right),$$
$$\widehat{\boldsymbol{\Psi}}=\sum_{i=1}^{m}E\left(\text{vec}(\mathbf{b}_i)\text{vec}(\mathbf{b}_i)^T|\mathbf{Q}_i,\mathbf{C}_i\right)/m, \quad (11)$$
$$\widehat{\Sigma}=\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\left(\text{vec}(\boldsymbol{\varepsilon}_{ij})\text{vec}(\boldsymbol{\varepsilon}_{ij})^T|\mathbf{Q}_i,\mathbf{C}_i\right)/\sum_{i=1}^{m}n_i.$$

The derivation of $\widehat{\Sigma}$ is similar to that of $\widehat{\boldsymbol{\Psi}}$ [25] since $\mathbf{b}_i$ is independent from $\boldsymbol{\varepsilon}_i$ and the rows of $\boldsymbol{\varepsilon}_i$ are independent of one another. The log-likelihood with the complete data has a similar functional form in terms of $\boldsymbol{\Psi}$ and Σ. We note the last equation in (11) differs from the approach in Hughes [8] in that a 2 by 2 matrix for Σ instead of a scalar is estimated, corresponding to a more general covariance structure for the measurement errors. The EM algorithm is stopped when the relative changes in the parameters are less than a tolerance value or maximum number of steps has been reached. The size of Markov chain is adjusted adaptively at each EM step by comparing the absolute changes in the estimates of fixed effect parameters to the Monte Carlo standard errors of the fixed-effects parameters [8].

Using Equations (3.1′) and (3.2′) in [26], the observed information on the censored data $\{(\mathbf{Q}_i,\mathbf{C}_i),1\le i\le m\}$ is equal to

$$\mathbf{I}_{Q,C}(\widehat{\boldsymbol{\theta}}_c)=\sum_{i=1}^{m}\left[E_{\widehat{\boldsymbol{\theta}}_c}\left(\mathbf{I}_{y_i}(\widehat{\boldsymbol{\theta}}_c)|\mathbf{Q}_i,\mathbf{C}_i\right)-E_{\widehat{\boldsymbol{\theta}}_c}\left(\mathbf{S}_{y_i}(\widehat{\boldsymbol{\theta}}_c)\mathbf{S}_{y_i}(\widehat{\boldsymbol{\theta}}_c)^T|\mathbf{Q}_i,\mathbf{C}_i\right)+\mathbf{D}_i\mathbf{D}_i^T\right], \quad (12)$$

where $\widehat{\boldsymbol{\theta}}_c$ is the maximum likelihood estimate from the MCEM method, $\mathbf{S}_{y_i}(\widehat{\boldsymbol{\theta}}_c)$ is the gradient of the log-likelihood on $\mathbf{y}_i$, $\mathbf{I}_{y_i}(\widehat{\boldsymbol{\theta}}_c)$ is the corresponding observed information matrix, $\mathbf{D}_i=E_{\widehat{\boldsymbol{\theta}}_c}(\mathbf{S}_{y_i}(\widehat{\boldsymbol{\theta}}_c)|\mathbf{Q}_i,\mathbf{C}_i)$. The derivations and formulas for the gradient, information

matrix, and the cross product of the gradient are given in the electronic appendix. To calculate $\mathrm{E}_{\hat{\theta}_c}(\mathbf{S}_{y_i}(\hat{\theta}_c)|\mathbf{Q}_i,\mathbf{C}_i)$ and $\mathrm{E}_{\hat{\theta}_c}(\mathbf{I}_{y_i}(\hat{\theta}_c)|\mathbf{Q}_i,\mathbf{C}_i)$, the Monte-Carlo expectations in Equation (10) evaluated at the MLE, $\hat{\theta}_c$, are used. To calculate $\mathrm{E}_{\hat{\theta}_c}(\mathbf{S}_{y_i}(\hat{\theta}_c)\mathbf{S}_{y_i}(\hat{\theta}_c)^T|\mathbf{Q}_i,\mathbf{C}_i)$, the following third and fourth-order conditional moments of $\mathbf{u}_i$ are also evaluated:

$$\mathrm{E}_{\hat{\theta}_c}\left(\mathbf{u}_i\,\mathrm{vec}(\mathbf{u}_i\mathbf{u}_i^T)^T|\mathbf{Q}_i,\mathbf{C}_i\right) \approx \sum_{l=1}^{L}\mathbf{u}_i^l\,\mathrm{vec}(\mathbf{u}_i^l(\mathbf{u}_i^l)^T)^T/L,$$

$$\mathrm{E}_{\hat{\theta}_c}\left(\mathrm{vec}(\mathbf{u}_i\mathbf{u}_i^T)\mathrm{vec}(\mathbf{u}_i\mathbf{u}_i^T)^T|\mathbf{Q}_i,\mathbf{C}_i\right) \approx \sum_{l=1}^{L}\mathrm{vec}(\mathbf{u}_i^l(\mathbf{u}_i^l)^T)\mathrm{vec}(\mathbf{u}_i^l(\mathbf{u}_i^l)^T)^T/L. \tag{13}$$

Inference based on the MCEM approach is similar to that for uncensored data (Equations (6), (7), (8)). The uncensored data estimate $\hat{\theta}_{uc}$, the gradient $\mathbf{S}_y(\hat{\theta}_{uc})$, and the information matrix $\mathbf{I}_y(\hat{\theta}_{uc})$ are simply replaced by the corresponding censored versions. Software implementing the MCEM algorithm in C and Fortran is available to readers upon request.

Note that the MI method has been modified by Fitzgerald *et al.* [27] to be in the spirit of an EM approach. Instead of drawing $\theta^{(t)}$ from the conditional distribution (Step 2 of MI), $\theta^{(t)}$ assumes the values of the MLE based on the current data while the other steps remain the same. This corresponds to a variation of the MCEM method in terms of parameter estimation. Consider the joint distribution $p(\mathbf{y}, \mathbf{b}, \mathbf{Q}, \mathbf{C}; \theta)$ where $\mathbf{y}$ is the uncensored data, $\mathbf{b}$ denotes the random effects, and $\mathbf{Q}$ and $\mathbf{C}$ are the censored data and censoring indicator respectively. The MCEM approach first determines the MLE for θ, assuming $\mathbf{y}$ and $\mathbf{b}$ are known, then takes the conditional expectation of the MLE estimate on $\mathbf{Q}$ and $\mathbf{C}$ at the current parameter estimate. The modified MI method finds the MLE for θ using the imputed data, then takes the expectation of the estimate over the imputed data sampled from $p(\mathbf{y}|\mathbf{Q}, \mathbf{C})$, which is in fact the Monte Carlo expectation. The advantage of the MCEM is that the information matrix for all the parameters can be used for making joint and robust inferences about the parameters [24].

3. Simulation studies

To compare the DL, HDL, MI, MLE, and MCEM methods, six sets of simulations under model (1) were conducted assuming parameter values consistent with the data from the example below and different censoring proportions and sample sizes. The model parameters were set to the following values:

$$(\beta_1, \beta_2)=(1.2, 2.0),$$

$$\begin{pmatrix} \sigma_{r11}^2 & \sigma_{r12} \\ \sigma_{r12} & \sigma_{r22}^2 \end{pmatrix} = \begin{pmatrix} 2.0 & 1.3 \\ 1.3 & 1.5 \end{pmatrix},$$

$$\begin{pmatrix} \sigma_{e11}^2 & \sigma_{e12} \\ \sigma_{e12} & \sigma_{e22}^2 \end{pmatrix} = \begin{pmatrix} 2.3 & 0.5 \\ 0.5 & 0.9 \end{pmatrix}.$$

It follows from these assumptions that ρ_r=0.75, ρ_e=0.35, and ρ=0.56. Three sets of censoring proportions for the first and second biomarkers were assumed: (0.20, 0.10), (0.35, 0.30) and (0.60, 0.50). For the number of subjects, we considered two scenarios: 300 subjects with 120 subjects having three repeated measures of each of the two biomarkers, and 180 subjects having four repeated measures; 100 subjects with 40 subjects having three repeated measures, and 60 subjects having four repeated measures. These values were chosen so that the simulated data would be similar to the HIV example described below. We first simulated values of $\mathbf{y}$, then the sample quantiles corresponding to the censoring levels for each biomarker were found; any simulated values below the corresponding quantiles for the

biomarkers were censored. Five hundred data sets were generated under each set of assumed conditions and the five methods were applied to each simulated data set. The DL and HDL methods were computed with R function *lme*. For the MI method, the following prior parameters were used: flat prior on the fixed effect parameters, inverse Wishart distribution with df = 2 and covariance parameter $0.5\mathbf{I}_2$ on both $\mathbf{\Psi}$ and $\mathbf{\Sigma}$. The Markov chain was started using the estimates from the HDL method with censored values imputed as HDL and then iterated 1800 steps. The values at step 1000, 1200, …, 1800 were used for multiple imputation. For both the MLE and MCEM approach, the initial values for the parameter θ were set to the corresponding estimates from the HDL method. The following settings were used in our computation with the MLE method, that is, the tolerance for quadrature of multivariate normal was 0.001; the three convergence values related to the sum of squares of changes in parameter estimates, the absolute change in log-likelihood, and the normalized gradients respectively were all set to 0.001; the maximal number of iterations was set to 200. The tolerance for the relative changes in parameters for the MCEM method was set to 0.001 for N=300 and censoring proportions = (0.20, 0.10) and (0.35, 0.30); to 0.005 for N=300 and censoring proportions = (0.60, 0.50), N=100 and censoring proportions = (0.20, 0.10) and (0.35, 0.30); to 0.01 for N=100 and censoring proportions = (0.60, 0.50). The burn-in steps of Markov chain was 150 and the starting size of Markov chain was 500. The maximal number of iterations for the MCEM method was set to 50.

The methods were compared with respect to mean relative bias, relative mean squared errors (relative MSE) and coverage probability of the 95% confidence interval [9, 28] for DL, HDL, MI, and MCEM. The references for mean relative bias and relative mean squared errors were both the true parameters. For simplicity, we'll call mean relative bias simply relative bias. We computed the empirical estimates of parameter estimates over the iterations for the MLE method and MCEM. Since the MLE method doesn't provide a confidence interval on the variance parameters and correlations, and more importantly the observed information matrices for the method were not positive definite in 68 (13.6%) simulations at N=300 and the highest censoring proportions, the ratios of model based standard error estimate (square root of mean variance estimates on parameters) to the empirical standard error (sample standard deviation of parameter estimates) were computed to assess the standard error estimates of the MLE method and MCEM. Only the simulations with positive definite information matrices were used when calculating the empirical mean and ratio of standard errors for the MLE method.

The results for N = 300 subjects are given in Table I and II. With a low censoring proportion: (0.20, 0.10), the relative biases of DL and HDL for $\beta_1,\beta_2, \sigma^2_{r11}, \sigma^2_{r22}, \sigma^2_{e11}, \rho$ are orders of magnitudes larger than the relative biases of MCEM, and are approximately two to three times greater than the relative biases for MI. The relative biases of DL and HDL in estimating the random effect correlation ρ_r are about 3 times the relative bias for MCEM and smaller than the relative bias for MI: –0.7%, –0.6%, –1.7%, and 0.2% for DL, HDL, MI, MCEM respectively. The relative MSE for the four methods: DL, HDL, MI, MCEM are more comparable, especially for the parameters $\beta_2, \sigma^2_{r22}, \rho_r, \rho_e, \rho$. The coverage probabilities with DL and HDL mostly differ from the target 95%, except for ρ_r and ρ_e. The coverage probabilities for ρ with DL, HDL, MI, and MCEM are 87.8%, 90.2%, 95.4%, 95.4% respectively. The coverage probabilities for all parameters with MCEM generally vary around 95% with a minimum coverage probability of 91.6% for β_1, and a maximum coverage probability of 96.2% for ρ_e. The coverage probabilities with MI have more variability, with a minimum of 50.4% for σ^2_{e11} and a maximum of 96.6% with ρ_e. In terms of empirical mean, the results from the MLE method and MCEM are almost identical which are both close to the true values, and the ratios of standard errors for the MLE and MCEM

methods are similar, both varying around 1. In this case, all the estimated information matrices for MLE method are positive definite.

As the censoring proportions increase, the DL, HDL and MI methods perform significantly worse. The performance of MCEM is also detrimentally affected, but it still remains within an acceptable range. For example, with DL and HDL, the relative biases for σ^2_{r11} increase from -34.6% and -28.4%, respectively, at the lowest censoring proportions to -80.7% and -75.7%, respectively at the highest censoring proportions. With MI, the relative bias for σ^2_{r11} increases from -13.2% to -38.7%; with MCEM the bias increases from -1.3% to -3.5%. The changes in relative biases and relative MSE for the other variance parameters, $\sigma^2_{r22}, \sigma^2_{e11}, \sigma^2_{e22}$, follow a similar pattern. The coverage probabilities of the four parameters for DL and HDL at the high censoring proportion are all 0; the coverage probabilities with MI are 5.0%, 35.4%, 0.6%, 2.0% for $\sigma^2_{r11}, \sigma^2_{r22}, \sigma^2_{e11}, \sigma^2_{e22}$ respectively, while they all remain around 95% with MCEM. The relative biases for ρ_r with DL, HDL, MI, MCEM at the high censoring proportions are -4.4%, -3.7%, -12.9%, 1.2% respectively, so that DL and HDL perform better than MI. The same is not true for ρ_e where MCEM performs better than MI, followed by HDL and DL. The order of performance for the coverage probabilities of ρ_r and ρ_e follows the same pattern. For the correlation parameter ρ, the relative biases for DL, HDL, MI at the low censoring proportions are -3.4%, -2.8%, -1.1% respectively; -15.9%, -14.0%, -11.6% at the high censoring proportions. In contrast, the relative bias for ρ with MCEM is 0.0% at the low censoring proportions and 0.1% at the high censoring proportions. The coverage probabilities for ρ with DL, HDL, MI, MCEM at the high censoring proportions are 15.8%, 21.4%, 59.4%, and 95.0% respectively. The empirical estimates of parameters for the MLE and MCEM methods are still similar at medium and high levels of censoring proportions. No simulation and 68 simulations out of 500 produce non-positive definite information matrices at medium and high levels of censoring, respectively, with the MLE method. The ratios of standard errors for the MLE method at high levels of censoring are similar to those for the MCEM method when restricted to the simulations with positive definite information matrices.

With a smaller sample size (number of subjects = 100), the relative biases of these four methods DL, HDL, MI, MCEM (Table III) remain similar to those observed with the larger sample size; the relative MSE of all four methods increases; the coverage probabilities for all parameters with DL, HDL, MI improve, but are still quite different from 95% for the variance parameters $\sigma^2_{r11}, \sigma^2_{r22}, \sigma^2_{e11}, \sigma^2_{e22}$ at the medium and high censoring proportions. The improvement in coverage probabilities can be attributed to the larger variance estimates for the parameters at the smaller sample size. The coverage probabilities for all parameter with MCEM remain at the 95% nominal level. The empirical estimates of parameters and the ratios of standard errors for the MLE and MCEM methods still follow the same pattern as in the N=300 case (Table IV). The ratios of standard errors for the MLE method are similar to the N=300 case; the numbers of non-positive definite information matrices at N=100 case were 1 (0.2%), 3 (0.6%), 62 (12.4%) for low, medium and high levels of censoring respectively.

Overall, the MCEM approach performs the best among DL, HDL, MI, MCEM in terms of relative bias, relative MSE, and coverage probability across all values of the censoring proportions except for one parameter, ρ_e. MCEM has negligible relative bias (<0.001) for ρ at low to medium levels of censoring, and only slightly overestimates β_1, ρ_r, and underestimates β_2 and σ^2_{r11}. The standard error estimates from the MCEM method are close to ideal. The performance of the MLE method is similar to that of the MCEM method in terms of parameter estimates as they both aim to obtain the MLE but using different numerical methods: one uses quadrature for likelihood and general optimization and the

other uses Monte-Carlo approximation to the EM algorithm. However, the MLE method is likely to result in non-positive definite information matrices particularly at high censoring proportions or when the sample size is small. The MI method performs better than the DL and HDL in terms of relative bias, relative MSE, and coverage probabilities at all levels of censoring proportions for all parameters except for ρ_r and σ_{e22}^2. For the parameter ρ_e, MI always performs slightly better than the MCEM in terms of relative MSE, which indicates that MCEM estimates ρ_e with more variability than MI. But MI overestimates β_1, σ_{r22}^2, σ_{e22}^2, ρ_e, and underestimates β_2, σ_{r11}^2, ρ_r, σ_{e11}^2, ρ. DL and HDL tend to overestimate the means, and underestimate the variances and all three correlation parameters. Underestimation of variation is a result of DL and HDL methods imputing fixed values for the censored observations, in contrast, the MI imputes random values and the MLE and MCEM methods find the MLE without imputing. The DL variance estimates are smaller than the HDL estimates since the former imputes the detection limits which are closer in magnitude to the non-censored values than half the detection limits. HDL consistently does a little better than DL with respect to relative bias, relative MSE, and coverage probability for all parameters.

4. Application

The data that we used to illustrate the methods is from the Women's Interagency HIV Study (WIHS), multi-center longitudinal study of HIV infection in US women. A random sample of 248 HIV sero-positive women was selected from the WIHS to study the correlation of HIV viral loads measured in the cervix and in plasma. The cervical HIV RNA levels and plasma HIV RNA levels were assessed at semi-annual visits, with the number of visits ranging from one to four (1.5 years of follow-up). The total number of person-visits was 1664. A significant proportion of HIV RNA levels was left censored: 57% of cervical HIV RNA and 26% of plasma HIV RNA values, with corresponding censoring values of 50 and 80 copies/mL, respectively. The proportions of person-visits with left censoring on both HIV RNA levels, cervical HIV RNA only, and plasma HIV RNA only were 23%, 34%, 3% respectively. The objective in this example is to assess the degree to which the cervical and plasma HIV RNA levels are correlated. We applied the DL, HDL, MI, MLE, MCEM methods to the log10 transformed HIV RNA data. The results are shown in Tables V and VI.

Generally the parameter estimates and confidence intervals from the DL and HDL are similar to each other but differ from the MI and MCEM results; the MI results differ from MCEM. The MLE method and MCEM have almost the same parameter estimates but their standard error estimates are slightly different. Specifically, the estimate of mean log10(cervical HIV RNA) is much higher with DL and HDL than with MI, MLE and MCEM. On the original scale for cervical HIV RNA, the estimates for the geometric mean HIV RNA level from DL and HDL are 288 (=10^2.46) and 194 (=10^2.29) copies/mL respectively; the estimates from MI, MLE and MCEM are 72, 28, 28 copies/mL respectively. Notice that 57% of cervical HIV RNA levels were below the detection limit of 50 copies/mL. The estimates from MI, MLE and MCEM therefore appear to be more plausible. The estimates for the geometric mean level of plasma HIV RNA from DL, HDL, MI, MLE, and MCEM methods are 1950, 1622, 912, 1148, 1148 copies/mL respectively.

As expected, the DL and HDL estimates of the variance parameters are smaller than those from the MI, MLE and MCEM methods. The MLE and MCEM methods yield higher variance estimates for σ_{r11}^2, σ_{e11}^2 (for cervical HIV RNA) than the MI and lower variance estimates for σ_{r22}^2, σ_{e22}^2 (for plasma HIV RNA) than the MI. The MI method is biased as shown in simulations. The direction of the bias depends on the underlying parameters and censoring proportions. Our simulations using parameter values similar to those in the

application showed that MI underestimates $\sigma^2_{r11}, \sigma^2_{e11}$ and overestimates $\sigma^2_{r22}, \sigma^2_{e22}$, so that inferences based on these parameters for MI may not be correct. We note that the DL and HDL correlation estimates for both the random effects and measurement errors are very similar. The estimate for ρ_r is higher with MCEM compared to the corresponding DL and HDL estimates, which in turn are higher than the MI estimate. The estimates for ρ_e with MCEM and MI are similar and both are higher than the DL and HDL estimates.

The final estimate and 95% confidence interval for ρ, the correlation between cervical HIV RNA and plasma HIV RNA, are 0.58 (95% CI: 0.53–0.63) with DL; 0.58 (95% CI: 0.52–0.63) with HDL; 0.57 (95% CI: 0.50–0.64) with MI and 0.66 (95% CI: 0.60–0.71) with MCEM. The MLE and MCEM methods have the same estimates of the correlation 0.66 and standard error estimates 0.03. The DL, HDL, MI estimates are all similar to each other and lower than the MCEM estimate, which is consistent with the patterns observed in our simulation studies.

5. Discussion

In this paper we proposed an MCEM method with a bivariate mixed-effects model for parameter estimation and inference in the presence of repeated measures and left-censoring. The performance of this method was shown in our simulations to be superior to the DL, HDL, and MI approaches with respect to relative bias, relative MSE, and coverage probability. Even at high censoring proportions, the MCEM still performs very well, while the DL, HDL, MI perform poorly. Although the MLE method can also estimate parameters accurately, our simulations showed that non-positive definite information matrices occur frequently, particularly at high levels of censoring, and therefore cannot serve as an alternative method for MCEM. In this paper, data from an HIV study was used to illustrate the proposed method but the approach naturally extends to other fields in which repeated measurements and censoring of biomarker levels are both present.

We used a relatively straightforward model and covariance structure in our derivations. Some possible extensions to the proposed approach include adjustment for covariates, more complex random effects structures, more complex measurement error covariance structures, and the case of more than two variables or biomarkers. To incorporate covariates into the model, we need only to replace $\mathbf{X}$ in model (1) with the specific design matrix which includes the covariates and modify β accordingly. To allow for more complex structures for the random effects, additional rows in the $\mathbf{b}_i$ matrix will be needed. Some examples include random effects for both the intercept and the slope, multi-level random effects, etc… One extension to the measurement error covariance structure can be achieved by assuming $\text{vec}(\varepsilon_i) \sim \text{N}(\mathbf{0}, \Sigma \otimes \mathbf{G}_i)$, where $\mathbf{G}_i$ is a n_i by n_i positive definite matrix with some defined structure: e.g., compound symmetry or AR(1) structure [11, 17]. Finally, one could consider more than two variables in the mixed-effects model and estimate the intra-class correlation. The derivation of the latter, however, will be more mathematically complex and implementation will also be more computationally challenging.

We have not addressed model misspecification with regard to distributional assumptions or the linear regression model assumption. Li *et al.* [19] considered the goodness-of-fit problem under censoring without repeated measures; the case involving repeated measures and censoring has not yet been addressed. The complex covariance structure for bivariate repeated measurements in the presence of censoring makes the assessment of model fit extremely difficult and is a potential area of future research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Bland JM, Altman DG. Correlation, regression, and repeated data. British Medical Journal. 1994; 308:896. [PubMed: 8173371]

2. Hamlett, A.; Ryan, L.; Wolfinger, R. On the use of PROC MIXED to estimate correlation in the presence of repeated measures. Proceedings of the Twenty-Ninth Annual SAS Users Group International Conference; SAS Institute Inc; 2004. Paper 198–29

3. Reichelderfer PS, Coombs RW, Wright DJ, Cohn J, Burns DN, Cu-Uvin S, Baron PA, Cohen MH, Landay AL, Beckner SK, Lewis SR, Kovacs AA. for the WHS 001 Study Team. Effect of menstrual cycle on HIV-1 levels in the peripheral blood and genital tract. AIDS. 2000; 14:2101–2107.10.1097/00002030-200009290-00005 [PubMed: 11061650]

4. Kovacs A, Wasserman SS, Burns D, Wright DJ, Cohn J, Landay A, Weber K, Cohen M, Levine A, Minkoff H, Miotti P, Palefsky J, Young M, Reichelderfer P. the DATRI and WIHS Study Groups. Determinants of HIV-1 shedding in the genital tract of women. Lancet. 2001; 358:1593–1601. [PubMed: 11716886]

5. Kemal KS, Foley B, Burger H, Anastos K, Minkoff H, Kitchen C, Philpott SM, Gao W, Robison E, Holman S, Dehner C, Beck S, Meyer WA III, Landay A, Kovacs A, Bremer J, Weiser B. HIV-1 in genital tract and plasma of women: compartmentalization of viral sequences, coreceptor usage, and glycosylation. Proceedings of the National Academy of Sciences USA. 2003; 100:12972–12977.10.1073/pnas.2134064100

6. Lynn HS. Maximum likelihood inference for left-censored HIV RNA data. Statistics in Medicine. 2001; 20:33–45.10.1002/1097-0258(20010115)20:1<33::AID-SIM640>3.0.CO;2-O [PubMed: 11135346]

7. Lau B, Gange SJ. Methods for the analysis of continuous biomarker assay data with increased sensitivity. Epidemiology. 2004; 15:724–732.10.1097/01.ede.0000142154.86749.e4 [PubMed: 15475722]

8. Hughes JP. Mixed effects models with censored data with application to HIV RNA levels. Biometrics. 1999; 55:625–629.10.1111/j.0006-341X.1999.00625.x [PubMed: 11318225]

9. 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; 1:355–368.10.1093/biostatistics/1.4.355 [PubMed: 12933561]

10. Chu H, Gange SJ, Li X, Hoover DR, Liu C, Chmiel JS, Jacobson LP. The effect of HAART on HIV RNA trajectory among treatment-naïve men and women: a segmental Bernoulli/Lognormal random effects model with left censoring. Epidemiology. 2010; 21:S25–S34.10.1097/EDE.0b013e3181ce9950 [PubMed: 20386106]

11. Roy A. Estimating correlation coefficient between two variables with repeated observations using mixed effects model. Biometrical Journal. 2006; 48:286–301.10.1002/bimj.200510192 [PubMed: 16708779]

12. Lyles RH, Williams JK, Chuachoowong R. Correlating two viral load assays with known detection limits. Biometrics. 2001; 57:1238–1244.10.1111/j.0006-341X.2001.01238.x [PubMed: 11764265]

13. Thiebaut R, Jacqmin-Gadda H, Babiker A, Commenges D. The CASCADE Collaboration. Joint modeling of bivariate longitudinal data with informative dropout and left-censoring, with application to the evolution CD4+ cell count and HIV RNA viral load in response to treatment of HIV infection. Statistics in Medicine. 2005; 24:65–82.10.1002/sim.1923 [PubMed: 15523706]

14. Albert PS. Modeling longitudinal biomarker data from multiple assays that have different known detection limits. Biometrics. 2008; 64:527–537.10.1111/j.1541-0420.2007.00886.x [PubMed: 17764483]

15. Hopke PK, Liu C, Rubin DB. Multiple imputation for multivariate data with missing and below-threshold measurements: time-series concentrations of pollutants in the Arctic. Biometrics. 2001; 57:22–33.10.1111/j.0006-341X.2001.00022.x [PubMed: 11252602]

16. Little, RJA.; Rubin, DB. Statistical Analysis with Missing Data. Wiley; New York: 1987.

17. Schafer JL, Yucel RM. Computational strategies for multivariate linear mixed-effects models with missing values. Journal of Computational and Graphical Statistics. 2002; 11:437–457.10.1198/106186002760180608

18. Pinheiro, JC.; Bates, DM. Mixed-effects Models in S and S-PLUS. Springer; New York: 2000. p. 79-82.

19. Li L, Wang WWB, Chan ISF. Correlation coefficient inference on censored bioassay data. Journal of Biopharmaceutical Statistics. 2005; 15:501–512.10.1081/BIP-200056552 [PubMed: 15920894]

20. Shao, J. Mathematical Statistics. 2. Vol. 61. Springer; New York: 2003.

21. Schafer JL. Multiple imputation: a primer. Statistical Methods in Medical Research. 1999; 8:3–15.10.1177/096228029900800102 [PubMed: 10347857]

22. Genz A. Numerical computation of multivariate normal probabilities. Journal of Computational and Graphical Statistics. 1992; 1:141–149.10.2307/1390838

23. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics. 1963; 11:431–441.10.1137/0111030

24. Boldea O, Magnus JR. Maximum likelihood estimation of the multivariate normal mixture model. Journal of the American Statistical Association. 2009; 104:1539–1549.10.1198/jasa.2009.tm08273

25. Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics. 1982; 38:963–974.10.2307/2529876 [PubMed: 7168798]

26. Louis TA. Finding the observed information matrix when using the EM algorithm. Journal of the Royal Statistical Society Series B. 1982; 44:226–233.

27. Fitzgerald AP, DeGruttola VG, Vaida F. Modeling HIV viral rebound using non-linear mixed effects models. Statistics in Medicine. 2002; 21:2093–2108.10.1002/sim.1155 [PubMed: 12111889]

28. Vaida F, Fitzgerald AP, DeGruttola V. Efficient hybrid EM for linear and nonlinear mixed effects models with censored response. Computational Statistics and Data Analysis. 2007; 51:5718–5730.10.1016/j.csda.2006.09.036 [PubMed: 19578533]

Appendix: Composition of WIHS HPV Study Group

WIHS HPV Study Group includes Andrea Kovacs (University of Southern California, Los Angeles, CA), Robert D. Burk (Albert Einstein College of Medicine, Bronx, NY), Howard Minkoff (Maimonides Medical Center and SUNY Downstate, Brooklyn, NY), Joel M. Palefsky (University of California at San Francisco, San Francisco, CA), Kathryn Anastos (Albert Einstein College of Medicine, Bronx, NY), Alexandra M. Levine (University of Southern California School of Medicine, Los Angeles, CA).

Table I

Comparison of four methods: DL, HDL, MI, MCEM in simulation studies (number of subjects = 300).

Cens. Prop.*	Criterion†	Method	β_1 (1.2)†‡	β_2 (2.0)	σ^2_{r11} (2.0)	σ^2_{r22} (1.5)	ρ_r (0.75)	σ^2_{e11} (2.3)	σ^2_{e22} (0.9)	ρ_e (0.35)	ρ (0.56)
(20%, 10%)	relBias (%)	DL	19.6	3.6	−34.6	−17.8	−0.7	−27.8	−14.0	−4.4	−3.4
		HDL	14.6	2.1	−28.4	−12.0	−0.6	−22.3	−8.9	−3.7	−2.8
		MI	6.2	−1.5	−13.2	7.4	−1.7	−10.1	9.0	2.4	−1.1
		MCEM	0.4	−0.1	−1.3	−0.3	0.2	0.0	0.0	0.3	0.0
	relMSE (%)	DL	4.6	0.3	12.7	3.9	0.3	7.9	2.2	1.0	0.4
		HDL	2.9	0.2	8.8	2.3	0.3	5.1	1.0	0.9	0.3
		MI	1.1	0.2	2.7	1.9	0.3	1.3	1.2	0.8	0.2
		MCEM	0.7	0.2	1.4	1.1	0.3	0.3	0.3	0.9	0.2
	CovProb (%)	DL	18.6	78.0	3.0	47.6	95.2	0	15.8	93.6	87.8
		HDL	41.6	89.2	15.6	71.8	95.2	0.4	52.2	94.0	90.2
		MI	82.2	94.2	74.6	86.8	95.0	50.4	69.4	96.6	95.4
		MCEM	91.6	93.4	94.8	94.6	95.6	95.6	95.4	96.2	95.4
(35%, 30%)	relBias (%)	DL	41.0	14.6	−54.9	−47.0	−1.9	−45.5	−38.0	−9.7	−7.5
		HDL	32.2	10.1	−48.0	−37.7	−1.5	−38.9	−29.2	−8.4	−6.3
		MI	12.4	−6.7	−21.8	27.2	−5.4	−16.7	30.9	5.2	−3.4
		MCEM	0.4	−0.1	−1.1	0.0	0.2	−0.1	0.1	0.4	0.0
	relMSE (%)	DL	17.6	2.3	30.6	22.5	0.4	20.8	14.6	1.9	0.9
		HDL	11.2	1.2	23.6	14.7	0.3	15.3	8.7	1.6	0.7
		MI	2.3	0.6	5.6	9.7	0.5	3.0	10.3	0.9	0.3
		MCEM	0.8	0.2	1.6	1.2	0.3	0.4	0.4	1.0	0.3
	CovProb (%)	DL	0	1.2	0	0	92.4	0	0	80.8	62.4
		HDL	1.2	16.8	0	0	94.6	0	0	84.8	72.4
		MI	56.8	70.2	46.0	50.6	88.4	15.8	7.2	97.4	93.2
		MCEM	93.0	92.8	93.6	95.0	96.4	95.2	94.4	94.8	95.8
(60%, 50%)	relBias (%)	DL	93.1	30.8	−80.7	−70.0	−4.4	−70.9	−59.3	−19.0	−15.9
		HDL	78.0	23.3	−75.7	−61.6	−3.7	−65.0	−50.1	−17.3	−14.0

Cens. Prop.*	Criterion†	Method	β_1 (1.2)†‡	β_2 (2.0)	σ^2_{r11} (2.0)	σ^2_{r22} (1.5)	ρ_r (0.75)	σ^2_{e11} (2.3)	σ^2_{e22} (0.9)	ρ_e (0.35)	ρ (0.56)
		MI	30.2	−12.8	−38.7	41.4	−12.9	−29.8	49.4	−1.5	−11.6
		MCEM	1.2	−0.2	−3.5	0.3	1.2	−0.4	0.2	−0.6	0.1
		DL	87.5	9.7	65.2	49.2	0.8	50.3	35.3	5.0	3.0
		HDL	61.8	5.6	57.4	38.2	0.7	42.3	25.2	4.2	2.4
	relMSE (%)	MI	10.1	1.9	15.7	20.2	2.0	9.1	25.5	0.6	1.5
		MCEM	1.0	0.2	2.3	1.5	0.4	0.7	0.5	1.4	0.3
		DL	0	0	0	0	81.6	0	0	46.6	15.8
		HDL	0	0	0	0	85.4	0	0	53.2	21.4
	CovProb (%)	MI	4.8	31.0	5.0	35.4	52.6	0.6	2.0	99.8	59.4
		MCEM	93.2	91.8	93.4	95.2	94.0	94.8	96.2	96.8	95.0

*
Cens. Prop. = censoring proportions (left-censoring) for biomarker 1 and 2 respectively.

†
relBias = mean relative bias with respect to the true parameter; relMSE = relative mean squared errors, defined as mean of (estimate − true parameter)^2/(true parameter)^2; CovProb = coverage probability of 95% confidence interval.

‡
The numbers in the parenthesis in the header row are the true parameter values.

Table II

Comparison of MLE method and MCEM in simulation studies (number of subjects = 300).

Cens. Prop.	Criterion[†]	Method	β_1 (1.2)	β_2 (2.0)	σ^2_{r11} (2.0)	σ^2_{r22} (1.5)	ρ_r (0.75)	σ^2_{e11} (2.3)	σ^2_{e22} (0.9)	ρ_e (0.35)	ρ (0.56)
(20%, 10%)	EmpEst	MLE	1.20	2.00	1.98	1.50	0.75	2.30	0.90	0.35	0.56
		MCEM	1.20	2.00	1.97	1.50	0.75	2.30	0.90	0.35	0.56
	RatioSE	MLE	0.92	0.95	1.01	1.00	0.93	1.05	1.03	1.02	0.93
		MCEM	0.92	0.95	0.96	0.95	0.98	1.03	1.01	1.03	0.98
(35%, 30%)	EmpEst	MLE	1.20	2.00	1.98	1.50	0.75	2.30	0.90	0.35	0.56
		MCEM	1.21	2.00	1.98	1.50	0.75	2.30	0.90	0.35	0.56
	RatioSE	MLE	0.93	0.94	1.03	1.02	0.95	1.05	1.04	1.03	0.94
		MCEM	0.92	0.94	0.97	0.96	0.99	1.02	1.02	1.04	0.98
(60%, 50%)	EmpEst	MLE[‡]	1.20	2.00	1.98	1.50	0.75	2.31	0.90	0.35	0.56
		MCEM	1.21	2.00	1.93	1.50	0.76	2.29	0.90	0.35	0.56
	RatioSE	MLE[‡]	0.98	0.94	1.10	1.09	0.97	1.08	1.09	1.07	0.98
		MCEM	0.94	0.92	0.94	1.00	0.93	0.98	1.05	1.07	1.00

[†] EmpEst = sample mean of parameter estimates from simulations; RatioSE = ratio of model based standard error estimates to empirical standard error, i.e., ratio of the square root of mean variance estimates for the parameters to the sample standard deviation of the parameter estimates.

[‡] The number of simulations out of 500 simulations with non-positive definite information matrices was 68 (13.6%). The results were based only on the simulations with positive definite information matrices.

Table III

Comparison of four methods: DL, HDL, MI, MCEM in simulation studies (number of subjects = 100).

Cens. Prop.	Criterion	Method	β_1 (1.2)	β_2 (2.0)	σ^2_{r11} (2.0)	σ^2_{r22} (1.5)	ρ_r (0.75)	σ^2_{e11} (2.3)	σ^2_{e22} (0.9)	ρ_e (0.35)	ρ (0.56)
	relBias (%)	DL	19.4	3.6	−34.3	−18.4	−0.3	−27.6	−14.6	−5.4	−3.5
		HDL	14.4	2.1	−28.1	−12.7	−0.2	−22.0	−9.5	−4.5	−2.9
		MI	5.4	−1.5	−11.8	6.2	−1.2	−8.9	8.0	2.2	−1.0
		MCEM	0.2	−0.1	−1.9	−1.0	0.8	0.5	−0.6	−0.6	−0.1
(20%, 10%)	relIMSE (%)	DL	5.6	0.6	13.8	5.6	0.8	8.1	2.7	3.2	0.7
		HDL	3.9	0.5	10.1	4.0	0.8	5.4	1.6	3.0	0.7
		MI	2.1	0.4	4.4	3.9	0.7	1.5	1.6	2.5	0.5
		MCEM	1.8	0.4	4.2	3.1	0.8	1.0	0.8	2.9	0.6
	CovProb (%)	DL	58.4	88.2	39.4	74.2	95.2	6.4	57.6	92.4	93.4
		HDL	73.0	92.4	57.6	84.4	95.6	18.8	77.8	94.0	94.0
		MI	90.8	95.4	91.0	95.0	96.6	83.2	92.8	95.8	97.6
		MCEM	93.4	96.2	93.8	95.2	96.2	94.4	95.2	94.8	95.6
	relBias (%)	DL	40.9	14.6	−54.8	−47.5	−1.2	−45.4	−38.4	−10.8	−7.6
		HDL	32.1	10.1	−47.9	−38.3	−0.8	−38.7	−29.5	−9.5	−6.4
		MI	12.3	−6.5	−21.8	24.7	−4.8	−16.8	30.5	5.1	−3.3
		MCEM	0.4	0.0	−2.0	−1.3	1.3	0.3	−0.4	−1.1	−0.1
(35%, 30%)	relIMSE (%)	DL	18.7	2.6	31.3	23.7	1.1	20.9	15.2	4.7	1.4
		HDL	12.3	1.5	24.4	16.1	1.0	15.4	9.2	4.2	1.1
		MI	3.5	0.9	7.3	12.2	1.0	3.6	11.4	2.6	0.6
		MCEM	2.0	0.4	4.7	3.5	0.9	1.3	1.1	3.6	0.7
	CovProb (%)	DL	5.4	21.8	2.2	3.4	94.8	0	0.2	84.8	84.2
		HDL	18.0	49.0	7.8	19.2	96.0	0.2	2.2	87.8	87.8
		MI	81.4	90.6	78.4	84.2	96.6	56.6	59.2	98.4	97.0
		MCEM	92.6	96.2	94.2	95.4	96.2	94.8	95.0	93.6	95.2
(60%, 50%)	relBias (%)	DL	93.0	30.6	−80.6	−70.1	−3.0	−70.8	−59.5	−20.0	−15.9
		HDL	78.0	23.1	−75.6	−61.8	−2.5	−65.0	−50.3	−18.4	−14.0

Cens. Prop.	Criterion	Method	β_1 (1.2)	β_2 (2.0)	σ^2_{r11} (2.0)	σ^2_{r22} (1.5)	ρ_r (0.75)	σ^2_{e11} (2.3)	σ^2_{e22} (0.9)	ρ_e (0.35)	ρ (0.56)
		MI	29.2	−13.4	−38.6	41.4	−12.9	−28.3	50.6	−1.6	−12.2
		MCEM	2.5	−0.3	−7.6	0.2	3.7	−0.3	0.1	−2.5	0.0
		DL	88.7	9.9	65.4	49.7	1.9	50.3	35.6	8.7	3.7
		HDL	63.0	5.8	57.7	38.9	1.6	42.4	25.6	7.7	3.0
	relMSE (%)	MI	11.0	2.4	17.0	25.9	2.6	8.8	30.0	2.2	2.1
		MCEM	2.8	0.5	7.5	4.4	1.5	2.3	1.7	5.7	0.9
		DL	0	0	0	0	90.2	0	0	70.0	53.6
		HDL	0	1.4	0	0	92.4	0	0	73.8	61.4
	CovProb (%)	MI	38.4	74.6	47.6	77.8	85.4	36.4	41.6	100.0	87.0
		MCEM	93.2	96.2	94.4	95.8	97.4	92.2	94.0	92.8	96.2

Table IV

Comparison of MLE method and MCEM in simulation studies (number of subjects = 100).

Cens. Prop.	Criterion	Method	β_1 (1.2)	β_2 (2.0)	σ^2_{r11} (2.0)	σ^2_{r22} (1.5)	ρ_r (0.75)	σ^2_{e11} (2.3)	σ^2_{e22} (0.9)	ρ_e (0.35)	ρ (0.56)
(20%, 10%)	EmpEst	MLE†	1.20	2.00	1.99	1.48	0.75	2.31	0.89	0.35	0.56
		MCEM	1.20	2.00	1.96	1.48	0.76	2.31	0.89	0.35	0.56
	RatioSE	MLE†	1.01	1.01	1.04	1.03	0.99	1.04	1.06	0.96	1.04
		MCEM	1.00	1.01	0.97	0.97	1.03	1.00	1.04	0.97	1.09
(35%, 30%)	EmpEst	MLE†	1.20	2.00	1.99	1.48	0.75	2.30	0.90	0.35	0.56
		MCEM	1.20	2.00	1.96	1.48	0.76	2.31	0.90	0.34	0.56
	RatioSE	MLE†	1.00	1.02	1.08	1.09	0.98	1.03	1.06	0.98	1.07
		MCEM	1.00	1.01	0.97	1.00	0.97	0.99	1.03	0.96	1.08
(60%, 50%)	EmpEst	MLE†	1.20	2.00	1.95	1.47	0.75	2.32	0.90	0.34	0.55
		MCEM	1.23	1.99	1.85	1.50	0.78	2.29	0.90	0.34	0.56
	RatioSE	MLE†	1.01	1.00	1.16	1.17	1.05	1.03	1.05	0.96	1.12
		MCEM	0.95	1.01	0.88	1.02	0.89	0.95	1.00	0.92	1.05

† The numbers of simulations out of 500 simulations with non-positive definite information matrices were 1 (0.2%), 3 (0.6%), 62 (12.4%) for censoring proportions (20%, 10%), (35%, 30%) and (60%, 50%) respectively. The results were based only on the simulations with positive definite information matrices.

Table V

Parameter estimates and 95% confidence intervals for the application data with DL, HDL, MI, and MCEM methods.

Parameter	DL	HDL	MI	MCEM
$\beta_{cervical}\ (=\beta_1)$	2.46 (2.36, 2.57)	2.29 (2.17, 2.41)	1.86 (1.67, 2.05)	1.45 (1.20, 1.70)
$\beta_{plasma}\ (=\beta_2)$	3.29 (3.17, 3.41)	3.21 (3.08, 3.34)	2.96 (2.77, 3.15)	3.06 (2.90, 3.23)
σ^2_{r11}	0.53 (0.41, 0.68)	0.66 (0.52, 0.85)	1.35 (1.03, 1.76)	2.22 (1.64, 3.00)
σ^2_{r22}	0.79 (0.64, 0.98)	0.93 (0.75, 1.15)	1.71 (1.32, 2.22)	1.41 (1.12, 1.79)
ρ_r	0.84 (0.76, 0.90)	0.84 (0.76, 0.90)	0.73 (0.63, 0.81)	0.88 (0.78, 0.93)
σ^2_{e11}	0.64 (0.57, 0.72)	0.80 (0.71, 0.89)	1.42 (1.20, 1.68)	2.29 (1.89, 2.77)
σ^2_{e22}	0.57 (0.51, 0.64)	0.67 (0.60, 0.75)	1.09 (0.92, 1.29)	0.87 (0.76, 1.00)
ρ_e	0.32 (0.24, 0.39)	0.31 (0.24, 0.39)	0.39 (0.27, 0.50)	0.40 (0.31, 0.49)
ρ	0.58 (0.53, 0.63)	0.58 (0.52, 0.63)	0.57 (0.50, 0.64)	0.66 (0.60, 0.71)

Table VI

Parameter estimates and standard errors (in parenthesis) for the application data with MLE and MCEM methods.

Method	β_1 (cervical)	β_2 (plasma)	σ_{r11}^2	σ_{r22}^2	ρ_r	σ_{e11}^2	σ_{e22}^2	ρ_e	ρ
MLE	1.45 (0.13)	3.06 (0.08)	2.23 (0.46)	1.41 (0.22)	0.87 (0.03)	2.29 (0.24)	0.87 (0.06)	0.41 (0.05)	0.66 (0.03)
MCEM	1.45 (0.13)	3.06 (0.08)	2.22 (0.34)	1.41 (0.17)	0.88 (0.04)	2.29 (0.22)	0.87 (0.06)	0.40 (0.05)	0.66 (0.03)