Hindawi
Journal of Probability and Statistics
Volume 2018, Article ID 1581979, 7 pages
https://doi.org/10.1155/2018/1581979

Review Article

Mixed Effects Models with Censored Covariates, with Applications in HIV/AIDS Studies

Lang Wu [1] and Hongbin Zhang [2]

[1] *Department of Statistics, University of British Columbia, Vancouver, Canada*
[2] *Department of Epidemiology and Biostatistics, Institute for Implementation Science in Population Health, City University of New York, USA*

Correspondence should be addressed to Lang Wu; lang@stat.ubc.ca

Received 26 March 2018; Accepted 5 May 2018; Published 3 June 2018

Academic Editor: Xu Zhang

Mixed effects models are widely used for modelling clustered data when there are large variations between clusters, since mixed effects models allow for cluster-specific inference. In some longitudinal studies such as HIV/AIDS studies, it is common that some time-varying covariates may be left or right censored due to detection limits, may be missing at times of interest, or may be measured with errors. To address these "incomplete data" problems, a common approach is to model the time-varying covariates based on observed covariate data and then use the fitted model to "predict" the censored or missing or mismeasured covariates. In this article, we provide a review of the common approaches for censored covariates in longitudinal and survival response models and advocate nonlinear mechanistic covariate models if such models are available.

1. Introduction

Mixed effects models are widely used in the analysis of clustered data, especially analysis of longitudinal data or survival data. In a longitudinal study, some variables are measured repeatedly over time, and these variables may be used either as responses or covariates, depending on study objectives. A common problem is that data on some of these variables may be left or right censored due to detection limits, may be missing at times of interest, or may be measured with errors. For example, in HIV/AIDS studies, viral load values may be left censored due to lower detection limits and may be missing or measured with substantial errors. In statistical analysis, these "incomplete data" issues must be addressed for correct statistical inference. In this article, we consider the case when these incompletely observed and time-varying variables are used as important covariates in mixed effects models for longitudinal response data or for time-to-event response data. To simplify the discussion, we focus on time-dependent covariates with left censoring, since similar methods/models may be used for right censoring or missing data or measurement errors in the covariates.

Longitudinal data with left censoring have received increasing attention in the literature in recent years (e.g., [1–8]). A common approach is to assume an empirical model for the covariate of interest based on the observed data, such as a linear mixed effects model. Then, the empirical model is used to "predict" the true covariate values when these values are censored, assuming the fitted model continues to hold for the unobserved censored values. A potential problem with this approach is that the assumed empirical covariate model based on the observed data may not hold for the censored covariate values, due to possibly different data-generation mechanisms for these "too small to observe" values. For example, in AIDS studies, censored viral loads below the detection limit may behave very differently from those above detection limit (observed values), due to a possibly different disease status for suppressed viral loads [6]. Moreover, the assumed model and distribution for the censored values cannot be verified based on observed data.

Recently, Kong and Nan [4] proposed an interesting approach based on ideas similar to that for right censored survival data, i.e., they used ideas similar to Cox models for right censored survival data for longitudinal data with left

censoring. Yu et al. [6] proposed an approach which treats censored values as point mass. While these two approaches make no distributional assumptions for the censored values, the methods may not be efficient if censored values indeed follow a parametric distribution similar to that for the observed values.

In some applications such as HIV viral dynamics and pharmacokinetic modelling, mechanistic or scientific models can be derived based on the underlying *data-generation mechanisms*. These models are often *nonlinear* and are derived based on a set of differential equations which approximately describe the true data-generation mechanisms, so these models are justified biologically or scientifically (e.g., [9, 10]). Moreover, these mechanistic models have been shown to fit observed data quite well based on many data analyses [11]. Since these mechanistic models are based on underlying true data-generation mechanisms, they should hold for *censored values*, even though these values are not observed. Therefore, these models can be used to better "predict" the unobserved censored values than empirical models. In this article, we will provide a review of such approaches. The approaches are illustrated by an HIV/AIDS dataset.

2. Mixed Effects Models with Censored Covariates

In this section, we focus on generalized linear mixed effects models for longitudinal responses and survival models for time-to-event responses, with left censored and time-dependent covariates. The methods can be extended to other types of regression models in a conceptually straightforward way.

2.1. Generalized Linear Mixed Models with Censored Covariates. We first consider generalized linear mixed models (GLMMs) with a left censored and time-dependent covariate in a longitudinal study, following Zhang et al. [7]. Let y_{ij} be the response of interest measured for individual i at time t_{ij}, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n_i$. Let x_{ij} be an important time-dependent covariate which is subject to left censoring, measurement errors, and missing data (assuming missing at random). We denote the unobserved true value of x_{ij} by x_{ij}^* in the presence of censoring or missing data or measurement errors. Let d be a known detection limit for x_{ij} such that x-values cannot be observed (detected) if $x_{ij} < d$ (i.e., left censoring), and let c_{ij} be the censoring indicator such that $c_{ij} = 1$ if $x_{ij} < d$ and $c_{ij} = 0$ otherwise. Let $\mathbf{z}_{ij}$ be a vector of other covariates.

Consider the following GLMM:

$$g\left(E\left(y_{ij}\right)\right) = x_{ij}^*\beta_1 + \mathbf{z}_{ij}^T\beta_2 + \mathbf{w}_{ij}^T\mathbf{a}_i, \quad \mathbf{a}_i \sim N(0, A), \; i = 1, \ldots, n; \; j = 1, \ldots, n_i, \tag{1}$$

where $g(\cdot)$ is a known link function, β_j's are unknown parameters, $\mathbf{w}_{ij}$ is a subset of $(x_{ij}, \mathbf{z}_{ij})$, $\mathbf{a}_i$ contains random effects, and A is a unknown covariance matrix. We assume that the response y_{ij} follows a distribution in the exponential family such as a normal or Poisson or Binomial distribution. When the covariate x_{ij} is left censored or missing or measured with error, we may assume an empirical model for x_{ij} based on the observed x-data, such as a linear mixed effects (LME) model. Then we assume that the LME model continues to hold for censored or unobserved values and proceed for likelihood inference. However, as noted in Section 1, such an approach may be problematic since censored values may not follow the same model obtained based on the observed data.

When a mechanistic or scientific model is available for covariate x_{ij}, such as in HIV viral dynamics, the scientific model should hold not only for observed data but also for unobserved data (e.g., censored or mismeasured or missing data), so that the model can be used to provide better "predictions" for the unobserved true covariate values. Such a scientific model is often *nonlinear*. For longitudinal data with large between-individual variations, by introducing random effects in the nonlinear model to account for between-individual variations and within-individual correlations among repeated measurements, we obtain a nonlinear mixed effects (NLME) model. Thus, we assume that the covariate x_{ij} follows the following NLME model:

$$x_{ij} = h\left(t_{ij}, \mathbf{b}_i, \alpha\right) + e_{ij} \quad \left(\equiv x_{ij}^* + e_{ij}\right), \quad \mathbf{b}_i \sim N(0, B), \; e_{ij} \text{ i.i.d.} \sim N\left(0, \sigma^2\right), \tag{2}$$

where $h(\cdot)$ is a known *nonlinear* function, vector $\mathbf{b}_i$ contains random effects, vector α contains fixed parameters, x_{ij}^* is the true covariate value at time t_{ij}, B is an unknown covariance matrix, and e_{ij}'s are random errors (measurement errors).

Note that when $h(\cdot)$ is a linear function (so model (2) is an LME model), the covariate model (2) is an empirical model which is chosen based on the observed covariate data. In a more general sense, the empirical models also include semiparametric or nonparametric mixed effects models. Such an empirical model is commonly used to address censoring, missing data, and measurement errors in the literature (e.g., [1, 2, 11]). When covariate x_{ij} is not normal, such as binary or count, generalized linear mixed models may be considered to fit observed covariate data, which are still empirical models. These empirical models may provide poor "predictions" to the unobserved data such as censored data.

2.2. Survival Models with Censored Time-Dependent Covariates. For survival models with time-dependent covariates, the covariates may also be left censored. Moreover, parameter estimation and inference for Cox models require that covariate values are available at event times [11]. However, this is usually not the case, since covariate values are unlikely to be available at all event times. Thus, this leads to missing covariate problems. The covariates may also be measured with errors, i.e., the observed covariate values may not be the true values but values with errors. In all cases, a common approach is to model the covariate process based on the observed covariate data and then use the fitted covariate model to "predict" the censored or missing covariate values. As noted in the previous section, a mechanistic or scientific covariate model may make better "predictions" than empirical covariate models, as shown in Zhang and Wu [8].

Here we consider a Cox model for the survival data with possible right censoring of the event times. For individual i, we define T_i to be the minimum of the observed event time T_i^* and the right censoring time C_i and define Δ_i to be the censoring indicator such that $\Delta_i = 1$ if the event time is right censored and $\Delta_i = 0$ otherwise, $i = 1, 2, \ldots, n$. Let $\lambda_i(t)$ be the hazard function for individual i at time t. The Cox model with time-dependent covariates can be written as

$$\lambda_i(t) = \lambda_0(t) \exp\left(\beta_1 x_i^*(t) + \boldsymbol{\beta}_2^T \mathbf{z}_i\right), \quad i = 1, 2, \ldots, n, \tag{3}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \beta_2)^T$ is a vector of regression coefficients and $\lambda_0(t)$ is the (unspecified) baseline hazard function.

When the time-dependent covariate $x_i(t)$ is left censored or missing or measured with errors, inference for the Cox model can be challenging. Similar to the GLMM in the previous section, a common approach is to model the time-dependent covariate $x_i(t)$ based on observed covariate data, assuming the fitted covariate model holds for the censored covariate values. Again, such an empirical approach can be problematic if censored covariate values behave quite differently than observed values. The problem can be fixed if a mechanistic covariate model is available. We may again consider the mechanistic NLME model (2) to address censoring in the covariates.

3. Statistical Inference

For parameter estimation and inference, two methods are commonly used: the two-step method and the joint likelihood method. We briefly review the two methods below.

3.1. Two-Step Methods. To estimate the parameters in the models, a simple approach would be the so-called *two-step method*: in the first step we fit the covariate model based on the observed covariate data, and then in the second step we fit the response model *separately*, with the censored or missing covariate values substituted by their predicted values from the first step.

Specifically, consider the GLMM response model (1) and the covariate model (2). In the first step, we fit the NLME covariate model (2) to the observed covariate data and obtain estimates of the parameters $\widehat{\alpha}$ and the empirical Bayes estimates of the random effects $\widehat{\mathbf{b}}_i$. The predicted value of the covariate at time t is given by

$$\widehat{x}_i(t) = h\left(t, \widehat{\mathbf{b}}_i, \widehat{\alpha}\right). \tag{4}$$

Then, in the second step, we fit the following GLMM to the response data using the standard complete-data method for fitting GLMM

$$g\left(E\left(y_{ij}\right)\right) = \widehat{x}_{ij}\beta_1 + \mathbf{z}_{ij}^T\beta_2 + \mathbf{w}_{ij}^T\mathbf{a}_i. \tag{5}$$

If the covariate x value is censored or missing or mismeasured at time t_{ij}, its value is imputed by the predicted value $\widehat{x}_i(t_{ij}) = \widehat{x}_{ij}$.

An obvious issue with the above simple two-step method is that the estimation *uncertainty* in the first step is ignored in the second step. The standard error of the parameter estimate $\widehat{\beta}_1$ may be underestimated, leading to misleading inference for the parameter β_1. To fix this problem, we may use the bootstrap method to obtain more reliable standard errors of the parameters in the response model [11]. A parametric bootstrap method, which generates samples from the above fitted models, may be used to produce more reliable standard errors of the estimates. Still, the two-step method may not be efficient because covariate data and response data are not used simultaneously.

If the response data are survival data, the issues mentioned above for the two-step method remain. Moreover, in this case, the longitudinal covariate data may be truncated by the events such as death or dropouts. In this case, the two-step method may lead to biased estimation.

3.2. Joint Likelihood Method. A more desirable and formal method than the two-step method is to use the likelihood method based on the "joint likelihood" for both the response and covariates. Maximum likelihood estimates (MLEs) of all the unknown parameters in the two models may then be obtained *simultaneously* based on the joint likelihood for all observed data. If all assumed models and distributions hold, the MLEs are the most efficient estimates. Let $\boldsymbol{\theta}$ be the collection of all unknown parameters in the response and covariate models, and let $f(\cdot)$ denote a generic density function. The joint log-likelihood for the observed data is given by

$$l_{obs}(\theta) = \sum_{i=1}^{n} \log \int \prod_{j=1}^{n_i} f_y\left(y_{ij} \mid x_{ij}, \mathbf{a}_i; \beta\right) f_x\left(x_{ij} \mid \mathbf{b}_i, \alpha\right)^{1-c_{ij}} F_x\left(d \mid \mathbf{b}_i, \alpha\right)^{c_{ij}} f\left(\mathbf{a}_i, \mathbf{b}_i\right) d\mathbf{a}_i d\mathbf{b}_i \tag{6}$$

where $f_y(y_{ij} \mid x_{ij}, \mathbf{a}_i; \beta)$ is a density function from the exponential family, $F_x(d \mid \mathbf{b}_i, \alpha) = P(x_{ij} < d \mid \mathbf{b}_i, \alpha)$, and c_{ij} is the censoring indicator for the covariates.

Evaluation of the intractable integration in the log-likelihood $l_{obs}(\theta)$ can be computationally challenging, especially when the dimension of the random effects $(\mathbf{a}_i, \mathbf{b}_i)$ is higher. By treating the random effects $(\mathbf{a}_i, \mathbf{b}_i)$ as "missing data," we may use the EM algorithm to find the MLEs. Let $\mathbf{x}_{i,cen}$ be the censoring components of the covariate vector $\mathbf{x}_i$. By treating $(\mathbf{a}_i, \mathbf{b}_i, \mathbf{x}_{i,cen})$ as "missing data", Zhang et al. [7] proposed a Monte Carlo EM algorithm in which the E-step is implemented with a Gibbs sampler combined with rejection sampling methods. The Monte Carlo EM algorithm is still computationally intensive but is feasible. Alternatively,

we may use computationally more efficient Laplace approximations or linearization methods to $l_{obs}(\theta)$ for approximate inference [11].

For the survival response models, the joint log-likelihood is given by

$$l^*_{obs}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log \int_{-\infty}^{\infty} \int_{-\infty}^{d} \left[f(T_i, \Delta_i \mid \mathbf{a}_i; \boldsymbol{\alpha}, \boldsymbol{\lambda}_0, \boldsymbol{\beta}) \times f_x\left(x_{ij} \mid \mathbf{b}_i, \alpha\right)^{1-c_{ij}} F_x\left(d \mid \mathbf{b}_i, \alpha\right)^{c_{ij}} \times f\left(\mathbf{a}_i; \mathbf{A}\right) \right] d\mathbf{a}_i, \quad (7)$$

where

$$\begin{aligned} &f(T_i, \Delta_i \mid \mathbf{a}_i; \boldsymbol{\alpha}, \boldsymbol{\lambda}_0, \boldsymbol{\beta}) \\ &\quad = \left(\lambda_i\left(T_i \mid X_i^*(T_i); \boldsymbol{\alpha}, \boldsymbol{\lambda}_0, \boldsymbol{\beta}\right)\right)^{\Delta_i} \times S\left(T_i \mid X_i^*(T_i)\right), \end{aligned} \quad (8)$$

with $S(t)$ the survival function defined as $S(t) = \exp(-\int_0^t \lambda(s)ds)$. Statistical inference can again be based on a Monte Carlo EM algorithm, although the computation can be more tedious due to the nonparametric baseline hazard in the Cox model.

4. Examples

In the following, we show two examples from an HIV/AIDS study. In the first example, we consider a Poisson generalized linear mixed model with censored covariates. In the second example, we consider a Cox survival model with censored covariates. In both examples, the time-dependent covariate is subject to left censoring and is modelled by a NLME model to address the censoring as well as missing data and measurement errors. The methods were implemented by Monte Carlo EM algorithms in R. R code is available upon request.

4.1. Generalized Linear Mixed Models with Censored Covariates. We consider an AIDS longitudinal dataset and study how viral load (VL) may relate to CD4 counts over time during an anti-HIV treatment. Viral loads usually have a lower detection limit so that viral load values below the limit cannot be observed, i.e., viral load may be left censored. Moreover, viral loads may be missing or measured with errors. As an illustration, we view CD4 count (y_{ij}) as the response and VL as a time-dependent covariate (x_{ij}), and we model the longitudinal CD4 counts as a Poisson GLMM:

$$\begin{aligned} \log\left(E\left(CD4_{ij}\right)\right) &= \beta_{0i} + \beta_{1i}t_{ij} + \beta_{2i}VL^*_{ij} + \beta_3 TR_i \\ &\quad + \beta_4 t_{ij} \times TR_i \end{aligned} \quad (9)$$

where $\beta_{ki} = \beta_k + a_{ki}$, $k = 0, 1, 2$, a_{ki}'s are random effects, and TR denotes a treatment indicator. Since VL may be left censored and may be measured with errors, we consider the following mechanistic NLME model which is justified biologically [9, 10]:

$$VL_{ij} = \log_{10}\left(\alpha_{1i}e^{-\alpha_{2i}t_{ij}} + \alpha_{3i}e^{-\alpha_{4i}t_{ij}}\right) + e_{ij} \equiv VL^*_{ij} + e_{ij}, \quad (10)$$

where $\alpha_{ki} = \alpha_k + b_{ki}$, $k = 1, 2, 3, 4$, b_{ki}'s are random effects, and viral load values VL_{ij} are $\log_{10}$-transformed. The random effects are assumed to follow multivariate normal distributions with mean 0 and unstructured covariance matrices. As a comparison, we also fit observed VL data based on an empirical LME model (ELM):

$$VL_{ij} = \alpha_{1i} + \alpha_{2i}t_{ij} + \alpha_{3i}t^2_{ij} + \alpha_{4i}t^3_{ij} + e_{ij}. \quad (11)$$

The unknown parameters (α_k, β_k) are estimated using a Monte Carlo EM algorithm as described in Zhang et al. [7].

Figure 1 shows the NLME and ELM models fit to the observed viral loads of two randomly selected subjects, where the times are rescaled to be in $[0, 1]$. It suggests different fitted curves from the two covariate models. In particular, the predicted lines based on the NLME model fit the uncensored viral loads quite well; and for the censored portion, the lines follow the mechanistic model and preserve an overall nonlinear trend. On the other hand, the empirical LME model renders noticeable deviation of the fitted lines from the uncensored viral loads and imposes a linear or quadratic curve for the censored viral loads. Such discrepancies between the covariate model fitting, particularly in the censored portion, induce different parameter estimates in the response model. Table 1 summarizes the parameter estimates of the response CD4 model, with the covariate VL being fitted based on the NLME and ELM models, respectively. As we can see, the results of the parameter estimates are different. For example, the estimate of β_2, which measures the association between CD4 and VL, is significant at 5% level based on the NLME covariate model but is not significant based on the ELM covariate model. The results based on the NLME model should be more reliable since it provides more reliable predictions for the censored viral loads, since the NLME model may make better predictions for the unobserved censored values than the ELM method as the NLME model is based on the underlying data-generation mechanism which is the same for both observed and unobserved covariate values. The higher the percentage of censored/missing values is, the better the NLME model performs. This is confirmed by the simulation study in Zhang et al. [7].

4.2. Survival Models with Censored Covariates. As another example, we consider the foregoing dataset again, but now we focus on the occurrences of the first CD4:CD8 decline. The objective here is to determine if and how the time to the first CD4:CD8 decline may be related to treatment and viral load. We consider the following Cox survival model for the time to first CD4:CD8 decline:

$$\lambda_i(t) = \lambda_0(t) \exp\left(\beta_1 TR_i + \beta_2 VL^*_i(t)\right). \quad (12)$$

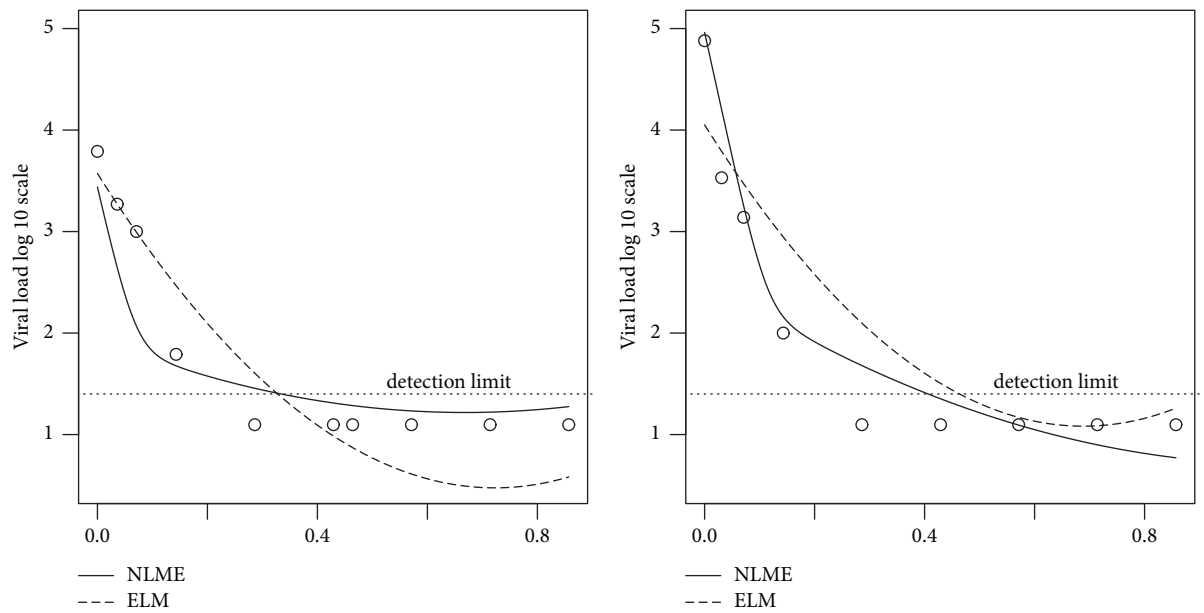

Figure 1: Fitted viral load curves for two randomly selected subjects. The open circles are the observed viral loads (the censored values are replaced by half the detection limit in $\log_{10}$ scale for simplicity). The solid line is the fitted curve based on the NLME model, while the dashed line is fitted curve based on the ELM model.

Table 1: Parameter estimates of the CD4 response model, based on the NLME and ELM covariate models respectively.

Response model parameter	NLME covariate model			ELM covariate model		
	Estimate	SE	p-value	Estimate	SE	p-value
β_0	5.98	0.16	0.00	6.15	0.23	0.00
β_1	−0.07	0.14	0.49	0.02	0.13	0.57
β_2	−0.19	0.06	0.01	−0.06	0.04	0.09
β_3	−0.71	0.59	0.19	−0.84	0.71	0.61
β_4	0.56	0.33	0.16	0.39	0.26	0.17

For this dataset, the Weibull distribution seems to provide a reasonable fit to the observed event times, so we consider the parametric Weibull distribution for the event times. For viral load, we use the same NLME and ELM models described in the first example.

Figure 2 shows, for two randomly selected subjects, the fitted lines to the observed viral loads based on the joint Cox survival model with the mechanistic NLME covariate model and empirical LME model (ELM), respectively, together with the corresponding estimated hazard functions and survival probability functions. We see that the mechanistic NLME model and the empirical LME model lead to different hazard and survival estimates. The NLME based joint model predicts monotonically increasing hazards, indicating the ever increasing risk of the event. On the other hand, the LME based model predicts more curved risk functions. Table 2 shows the results of the parameter estimates for the survival model. Here the differences seem relatively small, but as discussed, the predicted hazards and survival probabilities can be substantial. Since the NLME covariate model is derived based on reasonably biological justifications, they provide better "predictions" for censored (unobserved) viral loads and more reliable prediction for each individual's hazard and survival probability than the ELM covariate model, based on similar reasons as that for Table 1, which is also confirmed by simulations in Zhang and Wu [8].

5. Discussion

The nonlinear mechanistic covariate models are very appealing to address censoring and missing data in covariates, since the "predicted values" based on such models are more reliable than the commonly used empirical covariate models. These nonlinear mechanistic models are widely used in modelling HIV viral dynamics, pharmacokinetics, growth or decay, and some other areas [12, 13]. However, in many cases, such mechanistic models may not be available. In this case, an alternative approach is to treat the censored values as "point mass" to avoid unverifiable distributional assumptions for the censored values. The advantages of the nonlinear mechanistic covariate models are more obvious when the percentage of the censored values is higher, as confirmed in Zhang et al. [7]. The limitations of the nonlinear mechanistic covariate models are as follows: (i) in many

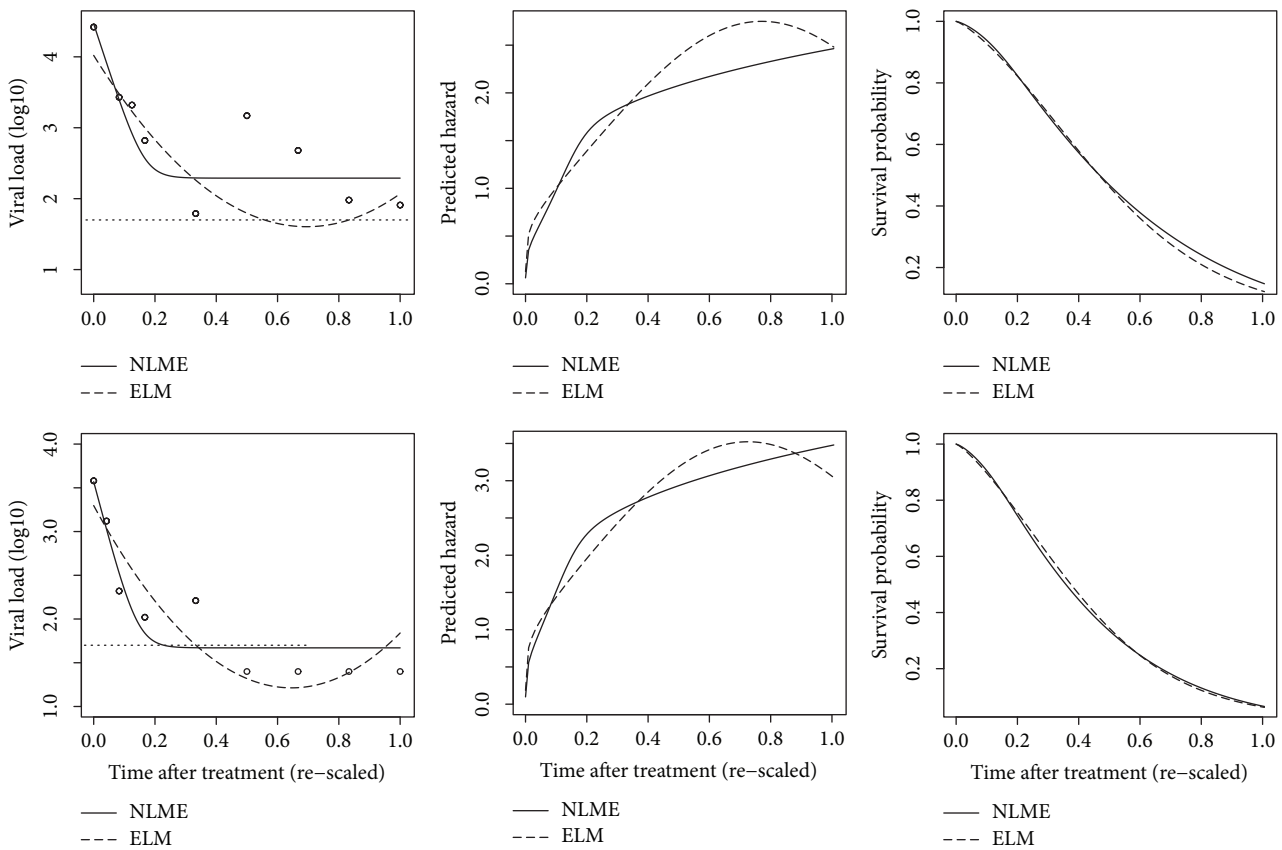

Figure 2: Plot of two individuals' (first row and 2nd row) fitted (predicted) viral load values and the corresponding hazard, survival functions based on the joint Cox and NLME model and empirical LME model (ELM), respectively. The open circles represent the observed viral loads. Left censored viral loads are replaced by one-half of the detection limit (in $\log_{10}$-scale).

Table 2: Parameter estimates in the Cox model based on the NLME and ELM covariate models respectively.

Cox model parameter	NLME covariate model			ELM covariate model		
	Estimate	SE	p-value	Estimate	SE	p-value
$\log(\lambda)$	1.60	1.05	0.10	1.44	1.13	0.60
$\log(\gamma)$	0.22	0.22	0.66	0.18	0.41	0.87
β_1	0.10	0.41	0.92	0.13	0.42	0.80
β_2	−0.40	0.54	0.39	−0.34	0.66	0.41

applications such mechanistic models may not be available and (ii) computation can be challenging, as discussed below.

Since the mechanistic covariate models are often nonlinear, computation is a main challenge in likelihood inference. Although Monte Carlo EM algorithms can almost always be used, they may offer potential problems such as very slow convergence or even nonconvergence. Moreover, the Monte Carlo EM algorithms usually need to be combined with Markov Chain Monte Carlo (MCMC) methods which are used to generate Monte Carlo samples in the E-step of the EM algorithms, making the computation even more challenging. When the dimensions of the random effects are high, we recommend approximate methods such as Laplace approximations and linearization methods as reviewed in Wu [11]. These approximate methods can be computationally much more efficient and provide reasonable approximations.

Data Availability

The dataset is available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is partially supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant no. 22R80742.

References

[1] P. W. Bernhardt, H. J. Wang, and D. Zhang, "Flexible modeling of survival data with covariates subject to detection limits via multiple imputation," *Computational Statistics & Data Analysis*, vol. 69, pp. 81–91, 2014.

[2] J. P. Hughes, "Mixed effects models with censored data with application to HIV RNA levels," *Biometrics*, vol. 55, no. 2, pp. 625–629, 1999.

[3] L. Wu, "A joint model for nonlinear mixed-effects models with censoring and covariants measured with error, with application to AIDS studies," *Journal of the American Statistical Association*, vol. 97, no. 460, pp. 955–964, 2002.

[4] S. Kong and B. Nan, "Semiparametric approach to regression with a covariate subject to a detection limit," *Biometrika*, vol. 103, no. 1, pp. 161–174, 2016.

[5] F. Vaida and L. Liu, "Fast implementation for normal mixed effects models with censored response," *Journal of Computational and Graphical Statistics*, vol. 18, no. 4, pp. 797–817, 2009.

[6] R. Fu and P. B. Gilbert, "Joint modeling of longitudinal and survival data with the Cox model and two-phase sampling," *Lifetime Data Analysis. An International Journal Devoted to Statistical Methods and Applications for Time-to-Event Data*, vol. 23, no. 1, pp. 136–159, 2017.

[7] H. Zhang, H. Wong, and L. Wu, "A mechanistic nonlinear model for censored and mismeasured covariates in longitudinal models, with application in AIDS studies," *Statistics in Medicine*, vol. 37, no. 1, pp. 167–178, 2018.

[8] H. Zhang and L. Wu, "A Mechanistic Nonlinear Model for Truncated and Mis-Measured Time-varying Covariates in Survival Models, with Applications in HIV/AIDS," *Journal of the Royal Statistical Society, C, accepted*, 2018.

[9] A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho, "HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time," *Science*, vol. 271, no. 5255, pp. 1582–1586, 1996.

[10] H. Wu and A. A. Ding, "Population HIV-1 dynamics in vivo: Applicable models and inferential tools for virological data from AIDS clinical trials," *Biometrics*, vol. 55, no. 2, pp. 410–418, 1999.

[11] L. Wu, *Mixed effects models for complex data*, vol. 113 of *Monographs on Statistics and Applied Probability*, CRC Press, Boca Raton, FL, 2010.

[12] J. K. Lindsey, *Nonlinear models in medical statistics*, vol. 26 of *Oxford Statistical Science Series*, Oxford University Press, Oxford, 2001.

[13] M. Davidian and D. M. Giltinan, *Nonlinear Models for Repeated Measurements Data*, Chapman Hall, London, 1995.