

RESEARCH

Open Access

Nonlinear mixed-effects HIV dynamic models with considering left-censored measurements

Tao Lu^{1*} and Min Wang²

*Correspondence:

stat.lu11@gmail.com

¹Department of Epidemiology and Biostatistics, State University of New York, Albany, New York, USA

Full list of author information is available at the end of the article

Abstract

HIV dynamic model offers a different perspective of studying HIV pathogenesis and developing treatment strategies for AIDS patients. Many HIV dynamic models have recently been developed to characterize short-term AIDS treatment, whereas in long-term HIV dynamics, viral load often rebounds in the later stage of treatment primarily due to reduced drug efficacy. Although time-varying drug efficacy can be incorporated into the ordinary differential equations (ODE) model, such a system has no analytical solution, and the measurement of viral load is usually censored at the detection limit due to technological constraints. We consider nonlinear mixed-effects ODE model with stochastic approximation EM algorithm to overcome these difficulties. The performance of the proposed method is illustrated by means of a simulation study and a real-data application. Numerical evidence shows that the HIV infection is generally more severe when considering left-censored data. The T cell production rate from human body source varies, but the death rate of infected T cells, infection rate of virus, and other dynamic parameters do not have much difference among patients. We hope these findings inspire more research on clarifying biological mechanism of HIV infection and developing better treatment.

Keywords: Below detection limit; Drug efficacy; HIV dynamic model; Long-term treatment; Nonlinear mixed-effect model

1 Introduction

HIV dynamic model, a set of ordinary differential equations (ODE) that describes the interaction between HIV virus and human body cells, has been proven useful for understanding the pathogenesis of HIV infection and developing treatment strategies. In order to estimate biologically/clinically meaningful parameters in the HIV dynamic model, many statistical models have been developed in the last decade, ranging from the simple nonlinear least squares (NLS) approach to more general nonlinear mixed-effect modeling approaches (Nowak and May 2000; Perelson and Nelson 1999; Perelson et al. 1996; Perelson et al. 1997; Wu and Ding 1999; Wu et al. 1998; Wu et al. 2005). However, in more complex scenarios, the ODE system has no close-form solution and thus needs to be solved numerically. Note that data from viral dynamic studies usually consist of repeated viral load measurements taken over time for each subject. In addition, the viral dynamic processes share certain similar patterns among patients while still having distinct individual characteristics. These properties indicate that the nonlinear mixed-effects (NLME) models appear to be reasonable for modeling HIV dynamics.

Disparate from short-term (a few hours) AIDS treatment, viral load remains sustained, whereas in long-term treatment (a few weeks to years), viral load usually decreases at the beginning of treatment and is followed by rebounds later, which leads to the treatment failure. There are various reasons that could contribute to the failure; one of which is the non-constant drug efficacy due to time-varying drug resistance, drug adherence or/and pharmacokinetical factors, such as median inhibitory concentration (IC_{50}), and so on. Huang et al. (2006) incorporated these clinical/PK factors into a time-varying drug efficacy model to characterize viral load resurgence featured in the long-term (24 weeks) AIDS treatment trajectory. This approach is capable of capturing the long-term treatment effect, but the difficulty may arise for the case in which the solution of ODE can not be linearized, because analytical formula does not exist. Accordingly, the classical nonlinear mixed-effects modeling approach becomes futile under this scenario.

It is remarkable that only a few methods have recently been proposed to tackle this problem in the longitudinal setting. For example, Huang et al. (2006) combined the Bayesian approach with the mixed-effects model to estimate both population and individual parameters within a framework of hierarchical Bayesian nonlinear models. From a frequentist point of view, Guedj et al. (2007) adapted a Newton-like approach using only first derivative to deal with the problem. Later on, in the presence of left-censoring, Huang and Getachew (2012) investigated the NLME model with a skew- t (ST) distribution for the response process and the ST nonparametric mixed-effects model for covariate measurement error process. However, the first approach ignores below detection limit viral load data but simply imputes the missing data with detection limit. The second approach requires reiteratively evaluating integral through an adaptive Gaussian quadrature approach, which quickly becomes intractable when the dimension of random effects is high. In addition, the computation of the approach is very complicated. The third approach is only applicable to short-term HIV dynamics and cannot be directly applied for long-term HIV dynamics. In particular, we may face severe bias in parameter estimation if we simply ignore censoring (below detection limit) or do not carefully handle it. As an illustration, in the simulation study performed later, we show that either imputing censored viral load with detection limit or completely removing those censored data will lead to large bias.

A classical way to deal with the incomplete or missing data problem is the Expectation-maximization (EM) algorithm, originally proposed by Dempster et al. (1977). For the E-step of the EM algorithm, one needs to compute the expectation of the complete log-likelihood with respect to missing data distribution at current estimates. The M-step includes the maximization of the expected log-likelihood function to update unknown parameters. One then follows iteration method until convergence is achieved. Within the mixed-effects modeling framework, the random effects can be treated as missing data so that the EM algorithm is capable of estimating fixed effects after averaging random effects in the E-step. Also, the below detection limit data are missing and can thus be averaged along with random effects at the E-step. Nevertheless, the E-step usually fails quickly when the number of censored data and/or the dimension of random effects increases since numerical integration becomes intractable in high dimensional situations.

To circumvent this difficulty, within linear mixed-effects model setting, Hughes (1999) used Monte Carlo integration to evaluate the E-step by repeatedly sampling from the

conditional distribution of below detection data. Wu (2002, 2004) extended the Monte Carlo version of EM (MCEM) for the nonlinear mixed-effects model. However, sampling the missing data is often time consuming and the algorithm is also slow to converge. Delyon et al. (1999) proposed a stochastic approximation version of EM (SAEM) to obtain the maximum likelihood estimate (MLE) of the unknown parameter. Kuhn and Lavielle (2005) coupled the Markov chain Monte Carlo (MCMC) procedure to the SAEM so that only one simulation of missing data at each iteration is required. The algorithm has been proven to converge quickly toward the MLE under general conditions. For theoretical proof, we refer readers to Delyon et al. (1999) and Kuhn and Lavielle (2005). In this paper, we couple the SAEM algorithm with the nonlinear mixed-effects model under the scenarios in which the ODE system has no analytical solution and the data from long-term AIDS treatment are left-censored. The performance of the proposed method is illustrated by means of a simulation study and a real-data application. Numerical evidence shows that the HIV infection is generally more severe when considering left-censored data.

The remainder of the paper is organized as follows. In Section 2, we introduce HIV dynamic model and time-varying drug efficacy model, which are used to model long-term AIDS treatment. In Section 3, the statistical model and method of estimating parameters are presented. In Section 4, we carry out a simulation study to examine the performance of the proposed method and in Section 5, we evaluate the performance of our method via analysis of the real AIDS clinical trial study. We conclude with a brief discussion in Section 6.

2 Model specification

2.1 Antiviral response model

One of the commonly used HIV dynamic models that describes the interaction between human body cells and virus is given by (Huang et al. 2006):

$$\begin{aligned} \frac{d}{dt}T(t) &= \lambda - \rho T(t) - [1 - \gamma(t)]kT(t)V(t), \\ \frac{d}{dt}T^*(t) &= [1 - \gamma(t)]kT(t)V(t) - \delta T^*(t), \\ \frac{d}{dt}V(t) &= N\delta T^*(t) - cV(t), \end{aligned} \tag{1}$$

where $T(t)$, $T^*(t)$, and $V(t)$ stand for uninfected T cells, infected T cells, and virus at time t , respectively, λ represents the rate at which new T cells are generated from body source like thymus, ρ is the death rate of uninfected T cells, k is the infection rate of T cells, δ is the death rate of infected T cells, N is the number of new virions produced from each infected T cells during their life-time, and c is the clearance rate of free virions. The time-varying parameter $\gamma(t)$ is the antiviral drug efficacy as defined below. In this paper, we assume that the system of equations (1) is in a steady-state before initiating antiretroviral treatment, and then the initial conditions for (1) are given by

$$T_0 = \frac{c}{kN}, \quad T_0^* = \frac{cV_0}{\delta N}, \quad \text{and} \quad V_0 = \frac{\lambda N}{c} - \frac{\rho}{k}. \tag{2}$$

The above ordinary differential equations have no analytical solutions and thus need to be solved numerically. One feature that distinguishes this dynamic model from others is

that by considering time-varying drug efficacy, we are capable of characterizing rebounds of virus in long-term antiviral treatment, which is not seen typically in the short-term period. On the other hand, more complexity may be added to the model by incorporating various clinical or/and pharmacokinetical (PK) factors such as drug susceptibility, drug resistance, and drug adherence etc. Thereafter, it becomes more challenging to carry out statistical inference.

2.2 Antiviral drug efficacy model

In many studies, drug efficacy is often assumed to be either a constant or perfect over treatment time; see, for example, Perelson et al. (1997) Perelson and Nelson (1999), Ding and Wu (2000), among others. Such an assumption may be reasonable during short-term period, whereas in long-term treatment, drug efficacy varies across time due to drug resistance, imperfect drug adherence or/and other clinical/pharmacological factors, which could lead to treatment failure. To deal with this issue, we advocate the modified E_{max} model to account for time-varying drug efficacy (Huang et al. 2006). In this paper, we consider two antiretroviral drugs given by

$$\gamma(t) = \frac{C_1 A_1(t)/IC_{50}^1(t) + C_2 A_2(t)/IC_{50}^2(t)}{\phi + C_1 A_1(t)/IC_{50}^1(t) + C_2 A_2(t)/IC_{50}^2(t)}, \quad (3)$$

where C_1 and C_2 represent drug concentration or any PK parameters such as C_{12h} . The median inhibitory concentrations $IC_{50}^1(t)$ and $IC_{50}^2(t)$ are used to quantify drug susceptibility (Molla et al. 1996) and are written as

$$IC_{50}(t) = \begin{cases} I_0 + \frac{I_r - I_0}{t_r} t & \text{for } 0 < t < t_r, \\ I_r & \text{for } t \geq t_r, \end{cases} \quad (4)$$

where I_0 and I_r are respective values of $IC_{50}(t)$ at the baseline and time point t_r , which is the time of virological failure. Let ϕ be a conversion factor between $IC_{50}(t)$ *in vitro* and $IC_{50}(t)$ *in vivo* and it can be estimated from data. Here, $A_1(t)$ and $A_2(t)$ are adherence profiles of the two drugs, which are measured by pill counts and are modeled as

$$A(t) = \begin{cases} 1 & \text{if all doses are taken in } (T_k, T_{k+1}], \\ R & \text{if } 100R\% \text{ doses are taken in } (T_k, T_{k+1}], \end{cases} \quad (5)$$

where T_k is the time at the k th clinical visit of patients, R_k represents the percentage of drugs that is consumed by patients during the interval of two visits. In particular, if all of the prescribed drugs are taken, R_k is one. This suggests that poor adherence is one of the major causes of treatment failure, and thus it is definitely necessary to take this factor into account; see, for example, Besch (1995).

3 Statistical model and parameter estimation

3.1 Nonlinear mixed-effects ODE model

Viral load data from an HIV/AIDS clinical trial are composed of repeated measurements on a group of patients so that a hierarchical modeling approach is necessary to account for within subject as well as between subject variation simultaneously. We are interested in estimating biologically/clinically meaningful parameters in (1) and conducting statistical inferences while taking below detection limit measurements for all patients into account. In the area of longitudinal data analysis, the mixed-effects model is often used to characterize both within and between subjects variation. Let y_{ij} be logarithmic

measurement of viral load for subject i at time t_{ij} for $i = 1, \dots, n$ and $j = 1, \dots, n_i$; population parameter $\boldsymbol{\mu} = (\log c, \log \delta, \log \lambda, \log \rho, \log N, \log k, \log \phi)^T$; individual parameter $\boldsymbol{\theta}_i = (\log c_i, \log \delta_i, \log \lambda_i, \log \rho_i, \log N_i, \log k_i, \log \phi_i)^T$. Let $g(\boldsymbol{\theta}_i, t_{ij}) = \log_{10}(V_{ij}(\boldsymbol{\theta}_i, t_{ij}))$ with $V_{ij}(\boldsymbol{\theta}_i, t_{ij})$ being the true amount of viral load based on (1) for subject i at time t_{ij} . Following Davidian and Giltinan (1995), a natural Nonlinear Mixed-effects ODE Model (NLME-ODE) is modeled as

- (i) Within-subject variation: $y_{ij} = g(\boldsymbol{\theta}_i, t_{ij}) + \boldsymbol{\varepsilon}_{ij}$. Measurement error $\boldsymbol{\varepsilon}_{ij}$ is assumed to follow a normal distribution with mean zero and variance σ^2 .
- (ii) Between-subject variation: $\boldsymbol{\theta}_i = \boldsymbol{\mu} + \mathbf{b}_i$. Random effect \mathbf{b}_i characterizes the deviation of individual parameters from population level and we assume $\mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$.

It deserves mentioning that the model described here is different from the classical NLME model in that $g(\cdot)$ has no explicit form in the long-term HIV dynamic setting, which leads to additional challenges for estimating parameters. A few methods such as the Bayesian approach and a Newton-like method were proposed to attack these challenges; see, for example, Huang et al. (2006), Guedj et al. (2007). However, these procedures either ignore missing (under detection limit) mechanism or are not applicable in the high dimensional setting. We here advocate the SAEM algorithm to take below detection limit data for parameter estimation into account in the longitudinal HIV dynamic model.

3.2 Parameter estimation

Both random effects from NLME-ODE and below detection limit (left-censoring) viral load can be treated as missing data. EM algorithm is able to deal with missing data where the log-likelihood with regard to missing components distribution is obtained at the E-step and the parameter estimates are updated through maximization (M-step). In light of high dimensionality of random effects and censored data, SAEM algorithm coupled with MCMC provides a convenient way of drawing samples at the E-step (Delyon et al. 1999; Kuhn and Lavielle 2005). The detail of this method is described below.

3.2.1 A. Expectation

From one point of view, both individual parameter $\boldsymbol{\theta}_i$ and below detection limit data can be treated as missing data. A classical way to cope with missing data is the EM algorithm proposed by Dempster et al. (1977). Let $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n)$. Denote \mathbf{Y}^o and \mathbf{Y}^m as observations beyond and below detection limit for all subjects across the study period, respectively. Let $L(\mathbf{Y}^o, \mathbf{Y}^m, \boldsymbol{\theta}; \boldsymbol{\mu}, \mathbf{D}, \sigma^2)$ represent the complete data likelihood. The MLE of $(\boldsymbol{\mu}, \mathbf{D}, \sigma^2)$ is determined by the marginal likelihood of the observed data $L(\mathbf{Y}^o; \boldsymbol{\mu}, \mathbf{D}, \sigma^2)$, whereas this quantity is often intractable. As an alternative, the EM algorithm calculates the expected value of the complete log-likelihood function, with respect to the joint conditional distribution of $\mathbf{Y}^m, \boldsymbol{\theta}$ given \mathbf{Y}^o under the current estimates of the parameter $(\boldsymbol{\mu}^{(k)}, \mathbf{D}^{(k)}, \sigma^{2(k)})$

$$\mathbf{Q}(\boldsymbol{\mu}, \mathbf{D}, \sigma^2 \mid \boldsymbol{\mu}^{(k)}, \mathbf{D}^{(k)}, \sigma^{2(k)}) = \mathbf{E}_{\mathbf{Y}^m, \boldsymbol{\theta} \mid \mathbf{Y}^o, \boldsymbol{\mu}^{(k)}, \mathbf{D}^{(k)}, \sigma^{2(k)}} [\log L(\boldsymbol{\mu}, \mathbf{D}, \sigma^2; \mathbf{Y}^o, \mathbf{Y}^m, \boldsymbol{\theta})]. \quad (6)$$

For simplicity of notation, let $I_o = \{(i, j); y_{ij} \geq \text{DL}\}$ with DL being the detection limit, and y_{ij}^o be the corresponding observation such that $I_m = \{(i, j); y_{ij} < \text{DL}\}$, and y_{ij}^m be the

corresponding missing data. Also, let n_t be the total number of observations and n_s be the number of subjects. In the long-term HIV treatment, it follows

$$L(\boldsymbol{\mu}, \mathbf{D}, \sigma^2; \mathbf{Y}^o, \mathbf{Y}^m, \boldsymbol{\theta}) \propto (\sigma^2)^{-\frac{n_t}{2}} \times |\mathbf{D}|^{-\frac{n_s}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i,j \in I_o} \left[y_{ij}^o - g(\boldsymbol{\theta}_i, t_{ij}) \right]^2 - \frac{1}{2\sigma^2} \sum_{i,j \in I_m} \left[y_{ij}^m - g(\boldsymbol{\theta}_i, t_{ij}) \right]^2 - \frac{1}{2} \sum_i (\boldsymbol{\theta}_i - \boldsymbol{\mu})^T \mathbf{D}^{-1} (\boldsymbol{\theta}_i - \boldsymbol{\mu}) \right\}. \tag{7}$$

Of particular note is that when $g(\cdot)$ is a linear function of $\boldsymbol{\theta}$, it is easy to obtain that equation (6) follows a normal distribution. For our problem, $g(\cdot)$ is the result of numerically integrating the ODE system in (1), which is not only a nonlinear function of $\boldsymbol{\theta}$, but the close-form expression does not exist as well. Accordingly, we follow the idea of a stochastic version of EM algorithm (SAEM) (Delyon et al. 1999) and evaluate equation (6) as follows.

3.2.2 B. Gibbs sampler for incomplete data

It has been long known that Gibbs sampler is useful for simulating data from the joint posterior distribution (Gelfand et al. 1990; Wakefield 1996). In our case, at the k th iteration, $\boldsymbol{\theta}$ and \mathbf{Y}^m can be alternatively generated from the joint posterior distribution $P(\boldsymbol{\theta}, \mathbf{Y}^m \mid \mathbf{Y}^o, \boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)}, \sigma^{2(k-1)})$ summarized in the following two steps.

Step 1 Simulate $\mathbf{Y}^{m(k)}$ from the marginal conditional posterior distribution $P(\mathbf{Y}^m \mid \boldsymbol{\theta}^{(k-1)}, \mathbf{Y}^o, \boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)}, \sigma^{2(k-1)})$ which follows a normal distribution truncated at the detection limit. Each $y_{ij}^{m(k)}$ is then centered at $g(\boldsymbol{\theta}_i^{(k-1)}, t_{ij})$ with variance $\sigma^{2(k-1)}$ and can thus be simulated as follows (Breslaw 1994):

- (a) calculate the cumulative probability of the detection limit under the same distribution as $y_{ij}^{m(k)}$ and denote as \mathbf{P}_{DL} ;
- (b) draw u from the uniform distribution $U(0, 1)$; and
- (c) obtain a sample of $y_{ij}^{m(k)}$ as $y_{ij}^{m(k)} = g(\boldsymbol{\theta}_i^{(k-1)}, t_{ij}) + \sigma^{(k-1)} \Phi^{-1}[u \times \mathbf{P}_{DL}]$, where Φ is the standard normal cumulative distribution function.

It should be noted that this sampling algorithm requires only one draw at each iteration therefore is efficient.

Step 2 Simulate $\boldsymbol{\theta}^k$ from the conditional posterior distribution $P(\boldsymbol{\theta} \mid \mathbf{Y}^{m(k)}, \mathbf{Y}^o, \boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)}, \sigma^{2(k-1)})$, which has no close-form formula but is proportional to (7) with all the parameters given at current values. The Metropolis-Hastings (M-H) algorithm is capable of generating samples from this distribution. Indeed, one choice of the proposal distribution is $q_{k-1} \sim \mathcal{N}(\boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)})$. Then the procedure proceeds as follows

- (a) Calculate acceptance probability $\alpha(\boldsymbol{\theta} \mid \boldsymbol{\mu}^{(k-1)})$ as

$$\min \left\{ 1, \frac{P(\boldsymbol{\theta} \mid \mathbf{Y}^{m(k)}, \mathbf{Y}^o, \boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)}, \sigma^{2(k-1)})}{P(\boldsymbol{\theta}^{(k-1)} \mid \mathbf{Y}^{m(k)}, \mathbf{Y}^o, \boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)}, \sigma^{2(k-1)})} \frac{q_{(k-1)}(\boldsymbol{\theta}^{(k-1)} \mid \boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)})}{q_{(k-1)}(\boldsymbol{\theta} \mid \boldsymbol{\mu}^{(k-1)}, \mathbf{D}^{(k-1)})} \right\}, \tag{8}$$

where φ is a candidate simulated from q_{k-1} . If we assume that each θ_i is independent, then \mathbf{D} is diagonal. We may thus simulate φ for each i (denote as φ_i) separately. After some arrangements, the acceptance probability α is simplified as

$$\alpha(\varphi^i | \theta^{(k-1)}) = \min \{1, \mathbf{R}^i\}, \tag{9}$$

where

$$\begin{aligned} \mathbf{R}^i = \exp \left\{ \frac{1}{2\sigma^{2(k-1)}} \left(\sum_{j;i,j \in I_o} [y_{ij}^o - g(\theta_i^{(k-1)}, t_{ij})]^2 + \sum_{j;i,j \in I_m} [y_{ij}^{m(k)} - g(\theta_i^{(k-1)}, t_{ij})]^2 \right. \right. \\ \left. \left. - \sum_{j;i,j \in I_o} [y_{ij}^o - g(\varphi_i, t_{ij})]^2 - \sum_{j;i,j \in I_m} [y_{ij}^{m(k)} - g(\varphi_i, t_{ij})]^2 \right) \right\}. \end{aligned} \tag{10}$$

- (b) For each i , draw u from the uniform distribution $U(0, 1)$. If $u \leq \alpha(\varphi^i | \theta^{(k-1)})$, then accept φ^i as new $\theta_i^{(k)}$; otherwise keep $\theta_i^{(k-1)}$ as $\theta_i^{(k)}$.

Notice that, unlike the implementation of M-H algorithm in other cases in which the choice of variance for proposal density q is essential to the efficiency of the algorithm. Herein, there is no need to choose an appropriate variance manually. Since variance \mathbf{D} is always estimated from the last iteration, the algorithm updates the proposal variance automatically to make itself adaptive. Often, the candidate parameter φ simulated from q makes the integration of (1) unstable, which is the so-called stiffness. To handle the stiff ODEs, we apply a Rosenbrock method, which is relatively easy to implement and also provides a good accuracy. We refer the interested readers to Kaps and Rentrop (1979) for more details.

3.2.3 C. Maximization

Once θ and \mathbf{Y}^m are simulated, it is straightforward to update $(\mu, \mathbf{D}, \sigma^2)$ by maximizing equation (6) and we obtain

$$\begin{aligned} \mu^{(k)} &= \frac{1}{n_s} \sum_i \theta_i^{(k)}, \\ \text{diag}(\mathbf{D}^{(k)}) &= \frac{1}{n_s} \sum_i \theta_i^{2(k)} - \mu^{2(k)}, \\ \sigma^{2(k)} &= \frac{1}{n_t} \left\{ \sum_{i,j \in I_o} [y_{ij}^o - g(\theta_i^{(k)}, t_{ij})]^2 + \sum_{i,j \in I_m} [y_{ij}^{m(k)} - g(\theta_i^{(k)}, t_{ij})]^2 \right\}. \end{aligned} \tag{11}$$

Note that the above estimates are composed of minimum sufficient statistics of $(\mu, \mathbf{D}, \sigma^2)$. Denote $s_1 = \sum_i \theta_i$, $s_2 = \sum_i \theta_i^2$, and $s_3 = \sum_{i,j \in I_o} [y_{ij}^o - g(\theta_i, t_{ij})]^2 + \sum_{i,j \in I_m} [y_{ij}^m - g(\theta_i, t_{ij})]^2$. Stochastic approximation step of SAEM is composed of updating s_1, s_2 , and s_3 with a sequence $\gamma^{(k)}$ at the k th iteration $s_i^{(k)} = (1 - \gamma^{(k)})s_i^{(k-1)} + \gamma^{(k)}s_i^{(k)}$. Kuhn and Lavielle (2005) recommended to use $\gamma^{(k)} = 1$ for the first $K1$ iterations followed by diminishing $\gamma^{(k)} = 1/(k - K1)$ for another $K2$ iterations in order to satisfy the assumptions of SAEM and to ensure the convergence of the algorithm. It deserves mentioning that the estimates of variance can be easily obtained by the inverse of the observed fisher information matrix of (7).

4 Simulation

We illustrate the effectiveness of our method via analysis of an example in the long-term HIV dynamic study with considering below detection limit measurements. Note that the simulation experiment is similar to an actual AIDS clinical trial. Basically, 40 patients were involved in a long-term antiretroviral treatment and then followed up to 200 days. Some patients were loss to follow up due to various reasons as in the real study. Consequently, the simulated viral load data are unbalanced longitudinal measurements. All data for the pharmacokinetic factor (C_1, C_2) , phenotype marker (baseline and failure IC_{50s}) and adherence as well as the baseline viral load (V_0) were taken from an AIDS clinical trial study.

In order to avoid parameter identifiability problem, ρ, N , and K are fixed at 0.1, 980, and 0.0001, respectively. These values are chosen from previous studies in the literature (Ding and Wu 2000; Nowak and May 2000; Perelson and Nelson 1999). In addition, we assume the system is at steady state at the beginning of the treatment. This indicates that only three parameters in (1) need to be estimated. Denote $\boldsymbol{\mu} = (\log c, \log \delta, \log \phi)^T$ and $\boldsymbol{\theta}_i = (\log c_i, \log \delta_i, \log \phi_i)^T$. The diagonal covariance matrix \mathbf{D} has a vector of diagonal elements $d^2 = (d_c^2, d_\delta^2, d_\phi^2)^T$. The value of $\log \lambda$ is obtained from (2). Altogether, there are totally seven parameters in the model.

The true values for $\boldsymbol{\mu} = (\log c, \log \delta, \log \phi)^T$ and $d^2 = (d_c^2, d_\delta^2, d_\phi^2)^T$ are set to be $(1.1, -1, 3)^T$ and $(0.25, 0.04, 0.0625)^T$, respectively. These values are similar to the ones used in the simulation study of Huang et al. (2006). The true value for σ^2 is 0.1 and the detection limit is 25 as in the real study. The true values of parameters $\boldsymbol{\theta}_i$ for $i = 1, \dots, 40$ are simulated from $\mathcal{N}(\boldsymbol{\mu}, \mathbf{D})$ independently. Given PK/clinical data $(C, IC_{50}, A(t))$ and true parameters $(\boldsymbol{\theta}_i)$ as well as initial viral load (V_0) , viral load V_{ij} are generated by numerically solving the ODE system in (1) for each individual. A measurement error term ε_{ij} is then added to $\log V_{ij}$ to mimic observation data $\log y_{ij}$ for the i th subject at time t_{ij} , where $\varepsilon_{ij} \sim N(0, \sigma^2)$. We then censor the viral load data at detection limit 25, producing totally 23 percent of below detection data which are imputed at 25 as observed in a real experiment.

We examine the effect of taking into account censored data or not on parameter estimates. First, the data below detection limit is completely hold out before fitting model, named as 'case A'; second, the below detection limit data are retained but assumed 25 as the true value, named as 'case B'; third, the below detection limit data are considered as censored data, named as 'case C'. For each case, we fit the nonlinear mixed-effects ODE model coupled with the SAEM algorithm. However, in cases A and B, the data are fit as they are, i.e., the data are considered as complete while censoring mechanisms are accounted for in case C. In this simulation study, K1 is set to be 10,000 to make sure the estimates converging to the neighborhood of MLE, then a decreasing step size γ is used for another 2,000 steps to fine-tune the final estimates. The same generating and model fitting processes are repeated 100 times. The performance of the proposed method on three cases is measured with respect to the bias and square root of MSE (RMSE).

The simulation results are presented in Table 1. In all three cases, the biases for inter-subject variance components are significantly larger than the ones of fixed effects. Biases for d_c^2 in all the three cases exceed 45%. The biases for variance estimates are especially large (almost 500% for 'case A' and 'case B'; 83% for 'case C'). For fixed effects, case A is superior to case B in terms of both bias and RMSE. Among all the three cases, case C

Table 1 The true values and mean estimates of population dynamic parameters with 100 replications as well as the corresponding bias, RMSE

		$\log c$	$\log \delta$	$\log \phi$	d_c^2	d_δ^2	d_ϕ^2	σ^2
Case A								
	True value	1.1	-1	3	0.25	0.04	0.0625	0.01
	Mean estimate	0.9	-0.8	3.07	0.13	0.06	0.11	0.05
	Bias(%)	15.2	23.7	2.4	47.3	46.4	69.5	484
	RMSE(%)	78	69	4.6	85	198	147	658
Case B								
	True value	1.1	-1	3	0.25	0.04	0.0625	0.01
	Mean estimate	1.9	-0.8	3.6	0.03	0.067	0.669	0.06
	Bias(%)	71	22	19	90	68	970	490
	RMSE(%)	141	58	23	105	125	1013	687
Case C								
	True value	1.1	-1	3	0.25	0.04	0.0625	0.01
	Mean estimate	1.2	-1.0	3.02	0.12	0.036	0.048	0.02
	Bias(%)	9.9	0.2	0.9	53	9.5	23.3	83
	RMSE(%)	54	5	3	75	39	211	783

In case A, the below detection limit data are hold out before fitting the model; In case B, the below detection limit data are retained but assumed 25 as the true value; In case C, the below detection limit data are considered as censored data and censoring mechanism is introduced by the proposed method.

has the least bias and RMSE for all of the seven parameters, thus performs the best. In general, imputing censored viral load with the detection limit (25) leads to larger bias for both fixed effects and random effects variance estimates. Our simulation results show that simply omitting censored observations (case A) is even better than retaining the censored values (case B) in terms of parameter estimating accuracy. Handling below detection limit data with the proposed method gives us the least bias in the current setting.

We also compared the fitting results with the Bayesian method and Gaussian quadrature method. We observed that for cases A and B, all the three methods provide similar estimates (Table 2). However, the Bayesian method fails to handle with below detection data so that it can not deal with case C. In addition, we also tried to fit case C with the Gaussian quadrature method. Unfortunately, this method does not converge in our case.

To test the robustness of parameter estimates with regard to initial values, we carry out a sensitivity analysis. Three sets of initial settings are examined. One sets the mean vector $\mu = (\log c, \log \delta, \log \phi)^T$ at the true values $(1.1, -1, 3)^T$, the other two set μ at the two extremes $(3, -0.5, 5)^T$ and $(0.2, -0.2, 0.6)^T$, separately. Variance components $d^2 = (d_c^2, d_\delta^2, d_\phi^2)^T$ are also increased to $(5, 5, 5)^T$ and $(10, 10, 10)^T$, respectively. We then

Table 2 Comparison of bias(%) on parameter estimates by three methods: Bayesian, Gaussian quadrature and NLME-SAEM methods

Case	Method	$\log c$	$\log \delta$	$\log \phi$	d_c^2	d_δ^2	d_ϕ^2	σ^2
A	Bayesian	16.1	24.2	2.1	49.6	47.7	65.1	543
	Gaussian quadrature	15.6	24	2.2	53.1	44.3	75.5	512
	NLME-SAEM	15.2	23.7	2.4	47.3	46.4	69.5	484
B	Bayesian	78	20	22	78	75	998	657
	Gaussian quadrature	72	25	16	94	61	932	586
	NLME-SAEM	71	22	19	90	68	970	490

In case A, the below detection limit data are hold out before fitting each model; In case B, the below detection limit data are retained but assumed 25 as the true value. The number of simulation run is 100.

estimate parameters using the same method as above. The first 10,000 steps are followed by additional 2,000 steps. Figure 1 depicts the graphical diagnostic of convergence of SAEM algorithm for the three cases. As shown clearly, after the first 10,000 steps, all three chains converge to the neighborhood of true values and 2,000 more steps make the estimate more precise. Based on these results, the different initial values give similar parameter estimates, therefore the method is quite robust. We may thus conclude that parameter estimates using the proposed method are insensitive to the choices of initial values.

5 A real AIDS clinical study

In this section, we apply the proposed method to a real AIDS clinical trial study. This study was a Phase I/II, randomized, open-label, 24-week comparative study of the PK, tolerability and antiretroviral effects of two regimens of indinavir (IDV), zidovudine (ZDV), zalcitabine (ZC), plus two nucleoside analogue reverse transcriptase inhibitors (NRTIs) on HIV-1-infected subjects failing protease inhibitor (PI) containing antiretroviral therapies. The 44 subjects were randomly assigned to the two treatment regimens, Arm A (IDV 800 mg q12h+RTV 200 mg q12 h), and Arm B (IDV 400 mg q12h+RTV 400 mg q12h). Study visits occurred at pre-entry, entry (within 14 days of pre-entry) and days 7, 14, 28, 56, 84, 112, 140 and 168 of follow-up. Plasma HIV RNA testing was conducted at each study visit. The detection limit is 25 cp/ml. Clinical assessment and laboratory parameters were performed at all visit weeks with the exception of week 1. Phenotypic determination (IC_{50}) of antiretroviral drug resistance was performed at baseline and at the time of virological failure. The PK parameters of IDV and RTV were determined using noncompartmental methods. Pill counts were measured to monitor adherence at each study visit from weeks 2 to 24. Of the 44 subjects, 42 subjects were included in this analysis; of the remaining two subjects, one was excluded from the analysis because the PK parameters were not obtained and the other was excluded because the phenotype assay could not be completed on this subject. Acosta et al. (2004) gives more detailed descriptions of this study.

The analysis here is to investigate the effect of missing data on the parameter estimation of HIV dynamic model (1). There are totally 14 percent (45 out of 328) of below detection observations in this real study. We compare the parameter estimate results based on the full data, with and without taking censoring mechanism into account. To avoid parameter identifiability problem, we fix $c = 3.0$ and $N = 980$; see Stafford et al. (2000). There are

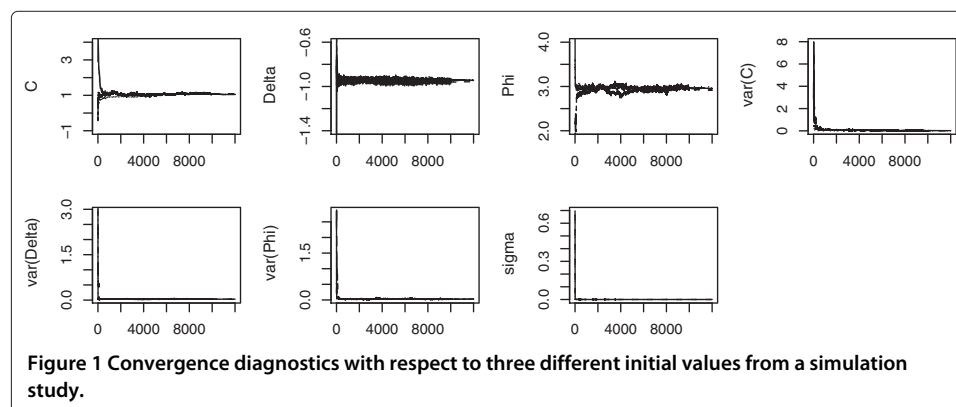


Figure 1 Convergence diagnostics with respect to three different initial values from a simulation study.

totally five fixed effects and six variance parameters (five from random effects and one from measurement error). We tried different initial values which all gave similar results. Table 3 presents parameter estimates and corresponding standard deviation for the 11 parameters.

The biological and clinical interpretation of estimating results are quite interesting. Compared with the estimate with considering censoring mechanism, the death rate of infected cell (T^*) is overestimated ($\log \delta = -1.613$ vs -1.801); the death rate of uninfected cell (T) is underestimated ($\log \rho = -3.798$ vs -3.563); the infection rate of T cells by virus is underestimated ($\log k = -11.87$ vs -9.397); conversion factor is overestimated ($\log \phi = 1.395$ vs 1.02); the production rate of T cells from body source is overestimated ($\log \lambda = 5.68$ vs 4.876) from the model without taking censoring into account. As a result, the comparison above consistently indicates that the degree of infection is underestimated by model without considering censoring if we define high degree of infection as low death rate of infected cell (T^*), high death rate of uninfected cell (T), high infection rate of T cells by virus, low production rate of T cells from body source. In other words, the degree of infection is actually more severe if censored data are taken care of. As one can see, these results are logical and meaningful. Moreover, as suggested by (3), conversion factor ϕ is negatively associated with drug efficacy. By considering censoring, the drug efficacy is higher, which is suggested by lower ϕ . Alternatively, the model without considering censoring is more likely to underestimate drug efficacy.

In addition, the estimates from the model that incorporates censoring mechanism are more precise in terms of smaller standard error except for $\log \lambda$. The inter-subject variance components are smaller for the ones based on the model with censoring mechanism, especially for k and ϕ . These results indicate most of parameter estimates are not much different among patients. On the contrary, the T cell production rate has a larger variation (d_λ^2) among patients when taking censoring into account. These results inform us that the dynamic parameters are not varying much from one subject to another, but each subject do present different health conditions as indexed by self production rate of T cells (λ). Again, these estimates are more precise compared to those from the model without considering censoring, as reflected by smaller standard errors. The variance for measurement error (σ^2) is not comparable since in the model with censoring, the censored data are replaced by viral load simulated from the model.

Table 3 Parameter estimates and standard deviation of parameters from a real AIDS clinical study using models with and without considering censoring mechanism, respectively

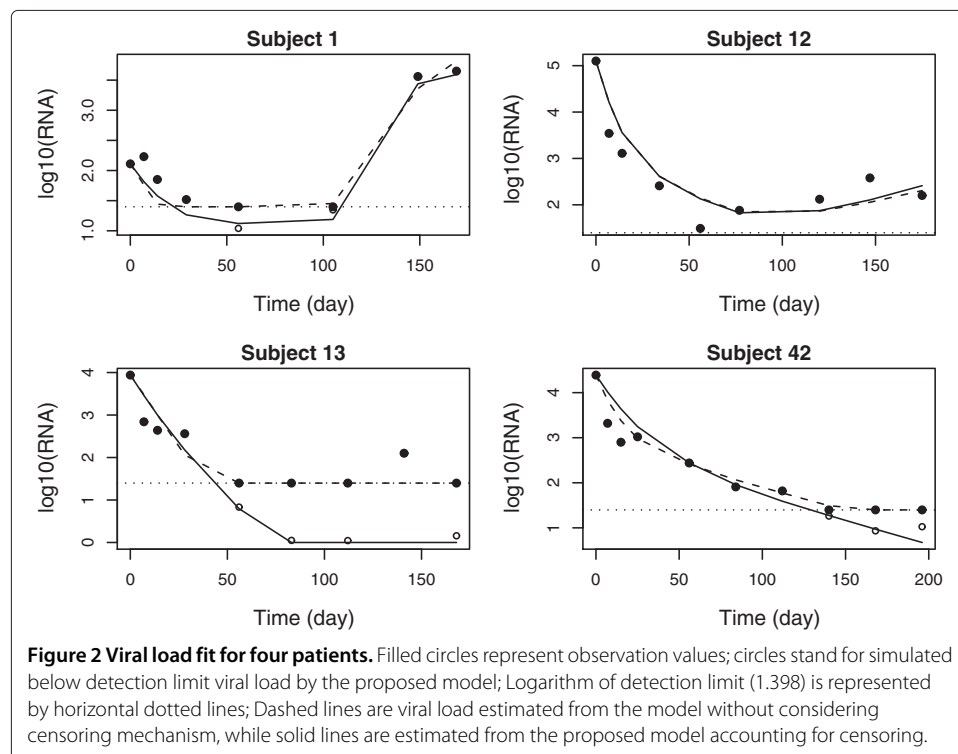
Parameter	W/o censor mechanism		With censor mechanism	
	Estimate	SE	Estimate	SE
$\log \delta$	-1.613	0.086	-1.801	0.07
$\log \rho$	-3.798	0.17	-3.564	0.02
$\log k$	-11.87	0.274	-9.397	0.001
$\log \phi$	1.395	0.094	1.02	0.001
$\log \lambda$	5.68	0.17	4.876	0.21
d_δ^2	0.312	0.068	0.203	0.044
d_ρ^2	1.219	0.266	0.016	0.004
d_k^2	3.146	0.686	7×10^{-4}	2×10^{-5}
d_ϕ^2	0.37	0.081	2×10^{-4}	3×10^{-6}
d_λ^2	1.215	0.265	1.787	0.39
σ^2	0.14	0.01	0.29	0.02

We also compared the model fitting from the two models with and without taking censoring for each patient into account. Individual parameter estimates are from the posterior mode of the Gibbs samples. The model fits the data very well. Figure 2 depicted the model fitting results of four patients. Clearly, the model not only describes consistently decreasing viral load trajectory (subject 13 and 42), but also rebounds (subject 1 and 12), which is a character of long-term treatment failure and a challenge to short-term models as explained in Section 1. When no below detection limit data are present, i.e. subject 12, the two models behave similarly. Even when censoring data occur, the two models fit the data beyond detection limit approximately the same way. However, without considering censoring, the model does not provide any information below detection limit. After accounting for censoring mechanism, we have a better idea of what is going on underneath the line. For example, in subject 1, viral load rebound has already occurred before it is detectable (solid line below detection limit). In subject 13, viral load keeps decreasing until diminished below detection limit.

Based on the analysis, we conclude that handling below detection data carefully is crucial to obtain more accurate estimates and also correct interpretation of the results.

6 Discussion

HIV dynamic model offers an opportunity to study HIV infection mechanism and treatment development from a perspective, other than laboratory experiments. However, research on this area is sparse mainly due to the absence of advanced statistical methods. The issue is further complicated by the lack of closed-form expression of the model and the numerical solution always incurs errors. In contrast to other methods introduced in Section 1, the SAEM algorithm is very efficient and fast, especially in the setting of



nonlinear mixed-effects ODE models. Numerical evidence shows that the algorithm converges very fast, usually in a few minutes whereas other algorithms, such as the Bayesian method, seem to be computationally expensive and also slow to converge. Note that the simulation step of SAEM employs a similar MCMC procedure, such as the Gibbs sampling and M-H step as in the Bayesian method. Nevertheless, the MCMC algorithm converges much faster when coupled with the SAEM as compared to the employment in a traditional Bayesian posterior sampling.

One disadvantage of the proposed method is that not all the parameters can be identified due to limitation of data on hand. We may use CD4 T cell counts in the model so that the constraints on some parameters can be released. However, the measurements of CD4 T cells are not reliable so that the use of these data may induce significant bias. The ODE model identifiability is an interesting topic but beyond the scope of this article. Some references for nonlinear ODE model identifiability, including HIV dynamic models, can be found in Jeffrey and Xia (2005), Wu et al. (2008), and Xia and Moog (2003).

In spite of this limitation, we compared the models with and without handling below detection data and found that ignoring these data is very dangerous for interpreting the results. The biological and clinical interpretation of the results accounting for censoring mechanism is inspiring. We hope the study results in this paper may stimulate more research in this area and contribute to the research on identifying the biological mechanism of HIV infection as well as clinical treatment development.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

TL initiated and carried out the study. TL drafted the manuscript. MW participated in the discussion and proofread the article. Both authors read and approved the final manuscript.

Acknowledgments

We would like to thank the editor and two anonymous referees whose valuable suggestions led to major improvements in the overall clarity and presentation of this paper.

Author details

¹Department of Epidemiology and Biostatistics, State University of New York, Albany, New York, USA. ²Department of Mathematical Sciences, Michigan Technological University, Houghton, Michigan, USA.

Received: 30 January 2014 Accepted: 14 March 2014

Published: 17 June 2014

References

- Acosta, EP, Wu, H, Hammer, SM, Yu, S, Kuritzkes, DR, Walawander, A, Eron, JJ, Fichtenbaum, CJ, Pettinelli, C, Neath, D, Ferguson, E, Saah, AJ, Gerber, JG: Comparison of two indinavir/ritonavir regimens in the treatment of HIV-infected individuals. *JAIDS*. **9**, 1358–1366 (2004)
- Besch, CL: Compliance in clinical trials. *AIDS*. **9**, 1–10 (1995)
- Breslaw, JA: Random sampling from a truncated multivariate normal distribution. *Appl. Math. Lett.* **7**, 1–6 (1994)
- Davidian, M, Giltinan, DM: *Nonlinear Models for Repeated Measurement Data*. Chapman & Hall, London (1995)
- Delyon, B, Lavielle, M, Moulines, E: Convergence of a stochastic approximation version of the EM algorithm. *Ann. Stat.* **27**, 94–128 (1999)
- Dempster, AP, Laird, NM, Rubin, DB: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B.* **39**, 1–38 (1977)
- Ding, AA, Wu, H: A comparison study of models and fitting procedures for biphasic viral decay rates in viral dynamic models. *Biometrics*. **56**, 16–23 (2000)
- Gelfand, AE, Hills, SE, Racine-Poon, A, Smith, AFM: Illustration of Bayesian inference in normal data models using Gibbs sampling. *J. Am. Stat. Assoc.* **85**, 972–985 (1990)
- Guedj, J, Thiebaut, R, Commenges, D: Maximum likelihood estimation in dynamical models of HIV. *Biometrics*. **63**, 1198–1206 (2007)
- Huang, Y, Getachew, D: Simultaneous Bayesian inference for longitudinal data with asymmetry, left-censoring and covariates measured with errors. *J. Japanes Stat. Soc.* **42**, 1–22 (2012)

- Huang, Y, Liu, D, Wu, H: Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics*. **62**, 413–423 (2006)
- Hughes, J: Mixed-effects models with censored data with applications to HIV RNA levels. *Biometrics*. **55**, 625–629 (1999)
- Jeffrey, AM, Xia, X: Identifiability of HIV/AIDS model. In: Tan, WY, Wu, H (eds.) *Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention*. World Scientific, Singapore, (2005)
- Kaps, P, Rentrop, P: Generalized Runge-Kutta methods of order four with stepsize control for stiff ordinary differential equations. *Numer. Math.* **33**, 55–68 (1979)
- Kuhn, E, Lavielle, M: Coupling a stochastic approximation version of EM with a MCMC procedure. *ESAIM: P. S.* **8**, 115–131 (2005)
- Molla, A, Korneyeva, M, Gao, Q, Vasavanonda, S, Schipper, P, Mo, H, Markowitz, M, Chernyavskiy, T, Niu, P, Lyons, N, Hsu, A, Granneman, G, Ho, DD, Boucher, C, Leonard, J, Norbeck, D, Kempf, D: Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nat. Med.* **2**, 760–766 (1996)
- Nowak, MA, May, RM: *Virus Dynamics: Mathematical Principles of Immunology and Virology*. Oxford University Press, Oxford (2000)
- Perelson, AS, Nelson, PW: Mathematical analysis of HIV-1 dynamics *in vivo*. *SIAM Rev.* **41**(1), 3–44 (1999)
- Perelson, AS, Neumann, AU, Markowitz, M, Leonard, JM, Ho, DD: HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span, and viral generation time. *Science*. **271**, 1582–1586 (1996)
- Perelson, AS, Essunger, P, Cao, Y, Vesanen, M, Hurley, A, Saksela, K, Markowitz, M, Ho, DD: Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*. **387**, 188–191 (1997)
- Stafford, MA, Corey, L, Cao, Y, Daar, ES, Ho, DD, Perelson, AS: Modeling plasma virus concentration during primary HIV infection. *J. Theor. Biol.* **203**, 285–301 (2000)
- Wakefield, JC: The Bayesian analysis to population Pharmacokinetic models. *J. Am. Stat. Assoc.* **91**, 62–75 (1996)
- Wu, L: A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies. *J. Am. Stat. Assoc.* **97**, 955–964 (2002)
- Wu, L: Exact and approximate inferences for nonlinear mixed-effects models with missing covariates. *J. Am. Stat. Assoc.* **99**, 700–709 (2004)
- Wu, H, Ding, AA: Population HIV-1 dynamics *in vivo*: applicable models and inferential tools for virological data from AIDS clinical trials. *Biometrics*. **55**, 410–418 (1999)
- Wu, H, Ding, AA, de Gruttola, V: Estimation of HIV dynamic parameters. *Stat. Med.* **17**, 2463–2485 (1998)
- Wu, H, Huang, Y, Acosta, EP, Rosenkaranz, S, Kuritzkes, D, Eron, JJ, Perelson, AS, Gerber, JG: Modeling long-term HIV dynamics and antiretroviral response: effects of drug potency, pharmacokinetics, adherence and drug resistance. *J. Acquir. Immune Defic. Syndr.* **39**, 272–283 (2005)
- Wu, H, Zhu, H, Miao, H, Perelson, AS: Parameter identifiability and estimation of HIV/AIDS dynamic models. *Bull. Math. Biol.* **70**, 785–799 (2008)
- Xia, X, Moog, CH: Identifiability of nonlinear systems with applications to HIV/AIDS models. *IEEE Trans. Automat. Contr.* **48**, 330–336 (2003)

doi:10.1186/2195-5832-1-13

Cite this article as: Lu and Wang: Nonlinear mixed-effects HIV dynamic models with considering left-censored measurements. *Journal of Statistical Distributions and Applications* 2014 **1**:13.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com