Hierarchical Bayesian Methods for Estimation of Parameters in a Longitudinal HIV Dynamic System

Yangxin Huang, Dacheng Liu and Hulin Wu
Department of Biostatistics & Computational Biology
University of Rochester, Rochester, NY 14642, USA

December 15, 2004

Corresponding Author:

Dr. Hulin Wu, Professor
Department of Biostatistics & Computational Biology
University of Rochester School of Medicine and Dentistry
601 Elmwood Avenue, Box 630, Rochester, NY 14642 USA
Phone: (585) 275–6767
Fax: (585) 273–1031
Email: hwu@bst.rechester.edu

Accepted in *Biometrics*

Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system

Yangxin Huang, Dacheng Liu and Hulin Wu
Department of Biostatistics & Computational Biology,
University of Rochester, Rochester, NY 14642, USA

SUMMARY

There have been substantial interests in investigating HIV dynamics for understanding the pathogenesis of HIV-1 infection and antiviral treatment strategies. Analysis of the dynamics of HIV-1 infection in response to drug therapy has elucidated crucial properties of viral dynamics. Although many HIV dynamic models have been proposed by AIDS researchers in the last decade, they have only been used to quantify short-term viral dynamics and are rarely applied to modeling virological response of long-term dynamics with the reduction in viral load either sustained or followed by a rebound, oscillation or a new viral load level. Since it is difficult to establish a relationship between antiviral response and pharmacokinetics due to many confounding factors related to antiviral responses, such as compliance and insurgence of resistance to treatment, the estimation of the dynamic parameters, therefore, has been mostly limited to the case of simplified, in particular, linearized models for a short-term viral dynamics. In this paper, a mechanism-based dynamic model is proposed for characterizing the long-term viral dynamics with antiretroviral therapy, described by a set of nonlinear differential equations without closed-form solutions. In this model we directly incorporate drug concentration, adherence and drug susceptibility into a function of treatment efficacy, defined as an inhibition rate of virus replication. We investigate a Bayesian approach under the framework of hierarchical Bayesian (mixed-effects) models for estimating unknown dynamic parameters. In particular, interest focuses on estimating individual dynamic parameters. The proposed methods can not only help to alleviate the difficulty in parameter identifiability, but also flexibly deal with sparse and unbalanced longitudinal data from individual subjects. For the purpose of illustration, we present two simulation examples to implement the proposed approach and apply the methodology to a real data set from an AIDS clinical trial. The basic concept of longitudinal dynamic systems and the proposed methodologies are generally applicable to any other biomedical dynamic systems.

KEY WORDS: Antiretroviral drug therapy; Bayesian mixed-effects models; drug exposures; drug resistance; HIV dynamics; MCMC; parameter estimation.

1. INTRODUCTION

Treatment of human immunodeficiency virus type 1 (HIV-1)-infected patients with highly active antiretroviral therapies (HAART), consisting of reverse transcriptase inhibitor (RTI) drugs and protease inhibitor (PI) drugs, results in several orders of magnitude of viral load reduction. The rapid decay in viral load can be observed in a relatively short term [45, 46, 66], and it either can be sustained or may be followed by a resurgence of virus within months [31]. The resurgence of virus may be caused by drug resistance, noncompliance, pharmacokinetics problems and other confounding factors during therapy. Mathematical models, describing the dynamics of HIV and its host cells, have been of essential importance in understanding the biological mechanisms of HIV infection, the pathogenesis of AIDS progression and the role of clinical factors in antiviral activities.

Many HIV dynamic models have been proposed by AIDS researchers [15, 25, 27, 40, 41, 44, 45, 46, 64, 66] in the last decade to provide theoretical principles in guiding the development of treatment strategies for HIV-infected patients, and have been used to quantify short-term dynamics. Unfortunately, these models are of limited utility in interpreting long-term HIV dynamic data from clinical trials. The main reason is that many parameters of these models can not be estimated uniquely from viral load data. It is due to this problem that simplified and linearized models have been often used to characterize the viral dynamics based on observed viral load data [25, 44, 45, 64, 65, 66]. Although these models are useful and convenient to quantify short-term viral dynamics, they can not characterize more complex long-term viral load trajectory. In this paper, we utilize a set of relatively simplified models, a system of differential equations with time-varying parameters, to characterize long-term viral dynamics. In our models we consider many factors related to the resurgence of viral load, such as the intrinsic nonlinear HIV dynamics, pharmacokinetics, compliance to treatment and drug susceptibility, and thus are flexible to quantify long-term HIV dynamics.

Data in viral dynamic studies usually consist of repeated viral load measurements on each of a number of subjects. Because the viral dynamic processes for different patients share certain similar patterns while still having distinct individual characteristics, the hierarchical nonlinear mixed-effects models appear to be reasonable for modeling HIV dynamics. Although mixed-effects modeling has been used in modeling pharmacokinetics/pharmacodynamics (PK/PD) for many years [54, 55], it was first applied to HIV studies by Wu *et al.* [65, 66]. Mixed-effects models offer a flexible framework where dynamic parameters for both individuals and population can be estimated by combining information across all subjects. The nonlinear mixed-effects (NLME) model fitting can be implemented in standard statistical software, such as the function `nlme()` in S-plus [47], the procedure `NLMIXED` in SAS [53] and other packages like NONMEM [6] and NLMEM [17]. In practice, the difficulty of using these standard packages in fitting NLME

models arises when the closed form of the nonlinear function is not available. For example, our viral dynamics nonlinear function is the solution of a system of nonlinear ordinary differential equations, which does not have a closed form. In this case the standard packages can not be used directly, and we must rely on the numerical solution to fit the mixed-effects models.

Bayesian statistics has made great strides in recent years. For various models, including the hierarchical NLME, parameter estimation and statistical inference are carried out via Markov chain Monte Carlo (MCMC) procedures [3, 4, 18, 19, 22, 23, 24, 35, 48, 49, 57, 62, 63]. The MCMC methods were introduced in NLME models with applications in PK/PD modeling in mid-1990s [3, 22, 35, 62, 63]. In particular, Bayesian analysis for a population HIV dynamic model was investigated by Putter *et al.* [48] and Han *et al.* [24] whose studies focused on the early dynamics of post-treatment HIV-RNA decline. Han *et al.* considered a linear dynamic model with the assumption that the number of uninfected target cells remained constant during a treatment. Under this assumption, the function of viral load over time t had a closed form. Although Putter *et al.* extended the model to a system of nonlinear differential equations, they only used short-term viral load data to estimate parameters, and they also assumed that the drug efficacy was constant. In addition, they did not consider the fact of variability in drug susceptibility (drug resistance) and adherence in the presence of antiretroviral therapy.

We combine Bayesian approach and mixed-effects modeling approach to estimate both population and individual dynamic parameters under a framework of the hierarchical Bayesian nonlinear (mixed-effects) model. In particular, we focus on estimating individual dynamic parameters by borrowing the population prior information to handle the identifiability problem at the individual level. The advantages of the proposed approach include: (i) we model the population viral dynamics, within-subject and between-subject variations via the hierarchical structure-based models; (ii) our models are flexible in dealing with both sparse and unbalanced longitudinal data from individual subjects by borrowing information from the population; (iii) the Markov chain Monte Carlo (MCMC) methods can be easily employed for computation, and thus closed-form solutions to the model, a system of nonlinear differential equations with time-varying coefficients, are not required; (iv) since the posterior distributions for the unknown parameters can be obtained, it is easy to conduct statistical inferences; and (v) the approach has more power to identify model parameters using the population prior information.

This paper is organized as follows. The concept of longitudinal dynamic systems is introduced in Section 2. We propose a simplified viral dynamic model with time-varying drug efficacy which incorporates the effects of pharmacokinetic (PK) variation, drug resistance and adherence in Section 3. A Bayesian approach implemented using MCMC techniques is employed to estimate dynamic parameters in Section 4. The simulation studies are used to assess and illustrate our methodologies in Section 5. The data for pharmacokinetics, drug resistance and adherence as well as the viral load data from an AIDS clinical trial are described and the proposed

methodology is applied to estimate the dynamic parameters in Section 6. Finally, the paper is concluded with some discussions in Section 7.

2. LONGITUDINAL DYNAMIC SYSTEMS

A dynamic system in engineering and physics, specified by a set of differential equations, is usually used to describe a dynamic process which follows physical laws or engineering principles. The idea of dynamic systems has been introduced into biomedical fields in recent years because of the rapid development of computational power and deep understanding of biological processes at cellular level with the assistance of modern biotechnologies. Examples of biological dynamic systems include genetic regulatory networks, tumor cell kinetics and viral dynamics.

Genetic regulatory networks can be described by a set of differential equations [10] as follow.

$$\begin{aligned} \tfrac{d}{dt}\mathbf{X} &= \boldsymbol{f}(\mathbf{Y}) - \boldsymbol{w}_1\mathbf{X}, \\ \tfrac{d}{dt}\mathbf{Y} &= \boldsymbol{w}_2\mathbf{X} - \boldsymbol{w}_3\mathbf{Y}, \end{aligned} \tag{1}$$

where the two differential equations represent two components: n-dimensional mRNA concentration vector ($\mathbf{X}$) and n-dimensional protein concentration vector ($\mathbf{Y}$) with n being the number of genes in the genome. $\boldsymbol{f}(\mathbf{Y})$ is n-dimensional translational function vector, and $n \times n$ non-degenerate diagonal matrices $\boldsymbol{w}_i$ $(i = 1, 2, 3)$ denote degradation rates of mRNAs, translational constants and degradation rates of proteins, respectively.

The models for tumor-immune system dynamics have been proposed by biomathematicians many years ago. For example, let X and Y denote two different kinds of tumor cells. Then a basic model for tumor cell kinetics, which is a set of differential equations, is given as [30, 37]

$$\begin{aligned} \tfrac{d}{dt}X &= r_1X(1 - K_1^{-1}X - w_1Y), \\ \tfrac{d}{dt}Y &= r_2Y(1 - K_2^{-1}Y - w_2X), \end{aligned} \tag{2}$$

where the r's are the growth parameters of the individual populations, the K's are the carrying capacities for the individual populations, and w's are interspecific competition rates. These models are successfully used in cancer research.

HIV dynamic models have been studied since late 1980s [25, 44, 45, 64, 66]. A simple model without antiviral treatment can be written [44]

$$\begin{aligned} \tfrac{d}{dt}X &= kTY - \delta X, \\ \tfrac{d}{dt}Y &= N\delta X - cY, \end{aligned} \tag{3}$$

where X is concentration of infected cells that produce virus, Y is free virus, T is target uninfected cells, k is the infection rate of T cells infected by virus, δ is the death rate for infected cells, N is the number of new virions produced from each of infected cells during their life-time, and c is the clearance rate of free virions. Similar models have been used for hepatitis virus

dynamics [42]. In this paper, we will focus our attention on studies of HIV dynamics, but models are more complicated than this simple model

The above models are general descriptions for genetic networks, tumor cell kinetics and viral dynamics for a single host. However, each individual host or subject may follow similar dynamic models but with different dynamic parameters. Thus, within-host and between-host variations need to be modeled appropriately. Some of the dynamic variables or their functions may be measured repeatedly for each individual in biomedical studies. This is similar to a longitudinal study but with dynamic systems. Thus, a longitudinal dynamic system can be described using a general framework of (hierarchical) nonlinear mixed-effects models (NLME) as follows. For individual i,

Stage 1. Within-subject variation:

$$\begin{array}{rcl} \frac{d}{dt}\boldsymbol{X}_i(t) & = & \boldsymbol{f}(\boldsymbol{X}_i(t), \boldsymbol{\theta}_i, t), \quad \boldsymbol{X}_i(0) = \boldsymbol{X}_{0i}, \ \ i = 1, \cdots, n, \\ \boldsymbol{Y}_i(t_j) & = & \boldsymbol{g}(\boldsymbol{X}_i(\boldsymbol{\theta}_i, t_j), \boldsymbol{\phi}_i, t_j) + \mathbf{e}_i(t_j), \quad \mathbf{e}_i|\boldsymbol{\theta}_i, \boldsymbol{\phi}_i \sim (\mathbf{0}, \boldsymbol{R}_i), \ \ j = 1, \cdots, m_i \end{array} \tag{4}$$

where $\boldsymbol{X}_i(t)$ is a vector of dynamic (latent) variables, $\boldsymbol{f}(\cdot)$ is known linear or nonlinear functions, $\boldsymbol{X}_{i0}$ is the initial values of the dynamic system, $\boldsymbol{Y}_i(t)$ is a vector of output or observation variables with a mean zero measurement noise $\boldsymbol{e}_i(t)$, $\boldsymbol{g}(\cdot)$ is a known observation function vector which can be linear or nonlinear, $\boldsymbol{\theta}_i$ is an unknown dynamic parameter vector and $\boldsymbol{\phi}_i$ is an unknown observation parameter vector.

Stage 2. Between-subject variation:

$$\begin{array}{rcl} \boldsymbol{\theta}_i & = & \boldsymbol{d}_1(\boldsymbol{\theta}, \boldsymbol{b}_{1i}), \quad \boldsymbol{b}_{1i} \sim (\mathbf{0}, \boldsymbol{D}_1), \\ \boldsymbol{\phi}_i & = & \boldsymbol{d}_2(\boldsymbol{\phi}, \boldsymbol{b}_{1i}), \quad \boldsymbol{b}_{2i} \sim (\mathbf{0}, \boldsymbol{D}_2), \end{array} \tag{5}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are the population parameter vectors, $\boldsymbol{b}_{1i}$ and $\boldsymbol{b}_{2i}$ are random vectors quantifying the between-subject variations.

The above mixed dynamic model is essentially a NLME model, but more complicated since a set of differential equations, which may not have a closed-form solution, is involved. Challenging problems for the longitudinal dynamic systems include identification of unknown dynamic parameters and statistical inference for dynamic parameters based on the data for the observed variables. In this paper, we introduce a longitudinal dynamic system for HIV dynamics, and suggest a Bayesian approach for dynamic parameter estimation and inference. Although we will concentrate on HIV dynamics in the following sections, the basic concept of longitudinal dynamic systems and the proposed methodologies are generally applicable to any other biomedical dynamic systems.

3. THE MODELS FOR LONG-TERM HIV DYNAMICS

3.1. Antiviral Drug Efficacy Model

Within the population of HIV virions in human hosts, there is likely to be genetic diversity and corresponding diversity in susceptibility to the various antiretroviral drugs. Recent treatment strategies usually include the genotype or phenotype tests in order to determine the susceptibility of antiretroviral agents before a treatment regimen is selected. Here we use the phenotype marker IC_{50} [38, 61], which represents the drug concentration necessary to inhibit viral replication by 50%, to quantify agent-specific drug sensitivity. Herein, we refer to this quantity as the median inhibitory concentration. To model the within-host changes over time in IC_{50} due to the emergence of new drug resistant mutations, we used the following function [27]

$$IC_{50}(t) = \begin{cases} I_0 + \frac{I_r - I_0}{t_r} t & \text{for } 0 < t < t_r, \\ I_r & \text{for } t \geq t_r, \end{cases} \tag{6}$$

where I_0 and I_r are respective values of $IC_{50}(t)$ at the baseline and time point t_r at which the resistant mutations dominate. Note that we use a time-varying parameter $IC_{50}(t)$ to model the susceptibility (resistance) of a virus population with a mixture of quasispecies resistant to the drugs in antiviral regimens. The make-up of these quasispecies may vary during treatment since drug resistant viral species may emerge if a sub-optimal dose of antiviral drugs is given [42]. Thus, separate equations for quasispecies are not necessary under our model setting. As examples, such functions for the ritonavir (RTV) and indinavir (IDV) drugs, respectively, are plotted in Figure 1(a). In clinical studies such as the one to be introduced in Section 6, it is common to measure IC_{50} values only at the baseline and failure time [38]. Then this simplified function of $IC_{50}(t)$ may serve as a good approximation.

Place Figure 1 here

Poor adherence to a treatment regimen is one of the major causes for treatment failure [5, 28]. Patients may occasionally miss doses, misunderstand prescription instructions or miss multiple consecutive doses for various reasons. These deviations from prescribed dosing affect drug exposure in predictable ways. For a time interval $T_k < t \leq T_{k+1}$, the effect of adherence can be defined as follows.

$$A(t) = \begin{cases} 1 & \text{if all doses are taken in } (T_k, T_{k+1}], \\ R & \text{if } 100R\% \text{ doses are taken in } (T_k, T_{k+1}], \end{cases} \tag{7}$$

where $0 \leq R < 1$ with R indicating adherence rate for drugs (in our clinical study, we focus on the two PI drugs of the prescribed regimen: RTV and IDV). T_k denotes the adherence evaluation time at the kth clinical visit. As an example, Figure 1(b) shows the effect of adherence over time for RTV and IDV drugs, respectively.

In recent years, antiretroviral drugs have been developed rapidly. The HAART has been proven to be extremely effective in reducing the amount of virus in the blood and tissues of infected patients. Three types of antiretroviral agents, nucleoside/non-nucleoside reverse transcriptase inhibitors (RTI) and protease inhibitors (PI) have been widely used to treat HIV infected patients in developed countries. Other types of antiviral agents, such as fusion inhibitors, are also in an active development stage. In most previous viral dynamic studies, investigators assumed that either drug efficacy was constant over treatment time [13, 16, 46, 66] or drugs had perfect effect in blocking viral replication [24, 25, 44, 45, 48] . However, the drug efficacy may change as concentrations of antiretroviral drugs and other factors (*e.g.* drug resistance) vary during treatment [13, 46], and thus the drugs may not be perfectly effective. Also in practice, patients' viral load may rebound during treatment and the rebound may be associated with resistance to antiretroviral therapy and other factors [15, 16]. To model the relationship of drug exposure and resistance with antiviral efficacy, we use a modified E_{max} model [56],

$$\gamma(t) = \frac{C_{12h}A(t)}{\phi IC_{50}(t) + C_{12h}A(t)} = \frac{IQ(t)A(t)}{\phi + IQ(t)A(t)}, \tag{8}$$

where $IQ(t) = C_{12h}/IC_{50}(t)$ denotes the inhibitory quotient (IQ)([26, 32]; the IQ is the PK adjusted phenotypic susceptibility and has recently been shown to be related to antiviral responses [26, 32]. $\gamma(t)$, ranged from 0 to 1, indicates the antiviral drug efficacy (the inhibition rate of viral replication) in a viral dynamic (response) model that will be introduced latter. C_{12h} is the drug concentration in plasma measured at 12 hours after a dose is taken. Note that C_{12h} could be replaced by other PK parameters such as the area under the plasma concentration-time curve (AUC) or the trough level of drug concentration (C_{min}). Although $IC_{50}(t)$ can be measured by phenotype assays *in vitro*, it may not be equivalent to the $IC_{50}(t)$ *in vivo*. Parameter ϕ indicates a conversion factor between $IC_{50}(t)$ *in vitro* and $IC_{50}(t)$ *in vivo*. This model is similar to the one used by Wahl and Nowak [60]. If $\gamma(t) = 1$, the drug is 100% effective, whereas if $\gamma(t) = 0$, the drug has no effect. Note that, if C_{12h}, $A(t)$ and $IC_{50}(t)$ can be observed from a clinical study and ϕ can be estimated from clinical data, then time-varying drug efficacy $\gamma(t)$ can be estimated during the course of antiretroviral treatment.

Model (8) quantifies the antiviral drug efficacy for a single drug from the RTI or PI class. To model the HIV dynamics under the HARRT, which usually contains three or more RTI/PI drugs, we need to extend model (8). For a regimen containing two agents within a class (for example, PI drugs), the combined drug efficacy of the two agents can be modeled as

$$\gamma(t) = \frac{IQ_1(t)A_1(t) + IQ_2(t)A_2(t)}{\phi + IQ_1(t)A_1(t) + IQ_2(t)A_2(t)}, \tag{9}$$

where $IQ_i(t) = C^i_{12h}/IC^i_{50}(t)$ $(i = 1, 2)$. Similarly, we can model the combined drug efficacy of an HAART regimen with both PI and RTI agents. Lack of adherence reduces the drug exposure,

which can be quantified by equation (7), and thus,based on the formula (8) or (9), reduces the drug efficacy which, in turn, can affect viral response. An example of the time course of the drug efficacy $\gamma(t)$ based on (9) with $\phi = 1$, $C^1_{12h} = 80$ and $C^2_{12h} = 50$ for two PI drugs is shown in Figure 1(c).

3.2. HIV Dynamic Model

Basic models of viral dynamics describe the interaction between cells susceptible to infection (target cells), infected cells and virus. The mathematical details of this model have been presented elsewhere [8, 27, 39, 42, 58]. In practice, we will need to trade-off between the model complexity and the identifiability of parameters based on available measurements from clinical trials. If a model has too many components, it may be difficult to analyze; many of the variables in the model may not be measurable and parameters may not be identifiable. If a model is too simple, although viral dynamic parameters can be identified and estimated, some important clinical factors, such as pharmacokinetics, adherence and drug resistance, and other biological mechanisms of HIV infection cannot be incorporated. A good model should be simple enough to incorporate available clinical data in the analysis, but detailed enough to reflect the major biological mechanisms and components in HIV infection. In order to consider clinical factors and biological mechanisms of HIV infection, and to flexibly deal with data analysis and identifiability of parameters, we propose an extended antiviral response model. Although our model only includes the interaction of target uninfected cells (T), infected cells that actively produce viruses (T^*) and free virus (V), this model differs from the previous models in that it includes a time-varying parameter $\gamma(t)$, which quantifies the antiviral drug efficacy. The model is expressed in terms of the following system of differential equations under the effect of an antiretroviral treatment

$$\begin{aligned} \tfrac{d}{dt}T &= \lambda - \rho T - [1-\gamma(t)]kTV, \\ \tfrac{d}{dt}T^* &= [1-\gamma(t)]kTV - \delta T^*, \\ \tfrac{d}{dt}V &= N\delta T^* - cV, \end{aligned} \tag{10}$$

where λ represents the rate at which new T cells are created from sources within the body, such as the thymus, ρ is the death rate of T cells, k is the infection rate of T cells infected by virus, δ is the death rate for infected cells, N is the number of new virions produced from each of infected cells during their life-time, and c is the clearance rate of free virions. The time-varying parameter $\gamma(t)$ denotes the antiviral drug efficacy as defined in the formula (8) or (9). If we assume that the system of equations (10) is in a steady-state before initiating antiretroviral treatment, then it is easy to show that the initial conditions for the system are

$$T_0 = \frac{c}{kN}, \quad T_0^* = \frac{cV_0}{\delta N}, \quad V_0 = \frac{\lambda N}{c} - \frac{\rho}{k} \tag{11}$$

If the regimen is not 100% effective (not perfect inhibition), the system of ordinary differential equations can not be solved analytically. The solutions to (10) then have to be evaluated numerically. Let $\boldsymbol{\beta} = (\phi, c, \delta, \lambda, \rho, N, k)^T$ denote a vector of parameters, and $\log_{10}(V_{ij}(\boldsymbol{\beta}, t))$ denote the common logarithm of the numerical solution of $V(t)$ for the ith individual at time t_j, which is the viral load measured in plasma and will be used for parameter estimation.

Similar to the analysis in [27], it can be shown from (10) that, in theory, if $\gamma(t) > e_c$ ($e_c = 1 - c\rho/kN\lambda$ is a threshold of drug efficacy) for all t, virus will be eventually eradicated. However, if $\gamma(t) < e_c$ (treatment is not potent enough) or if the potency falls below e_c before the viral eradication (due to drug resistance, for example), viral load may rebound (see [27] for a detailed discussion). Thus, the efficacy threshold e_c may reflect ability of a patient's immune system to control viral replications, and it is important to estimate e_c for each patient based on clinical data.

4. BAYESIAN APPROACH FOR PARAMETER ESTIMATION

In order to apply the proposed mathematical models to study long-term HIV dynamics and model viral responses, we need to resolve two important statistical problems: (i) how to estimate the unknown parameters in HIV dynamic models, and (ii) how to conduct inference and handle the identifiability of model parameters. It is challenging to achieve these two goals for a system of nonlinear differential equations with time-varying parameters, because there is no closed-form solution, and there are too many unknown parameters. In addition, among the components involved in viral dynamics we usually only have viral load data, where as the $CD4^+$ T cell count data are considered too noisy to be used for dynamic parameter estimation. It is possible that we may not be able to identify all the unknown parameters in the model (10). To deal with the identifiability problem of parameter estimation, mathematicians usually substitute some of the unknown parameters with their estimates from previous studies (see Perelson *et al.* [45] for example). Here we investigate a Bayesian approach to tackle this difficulty. In Bayesian terminology, the information from previous studies is regarded as prior knowledge which is combined with clinical data to perform the statistical inference on unknown parameters. A detailed description of the methodologies is given below.

4.1. Hierarchical Bayesian Modeling Approach

Under the longitudinal dynamic system framework, the hierarchical Bayesian approach can be used to incorporate a prior at the population level to estimate the dynamic parameters. We denote the number of subjects by n and the number of measurements on the ith subject by m_i. For notational convenience, let $\boldsymbol{\mu} = (\log\phi, \log c, \log\delta, \log\lambda, \log\rho, \log N, \log k)^T$, $\boldsymbol{\theta}_i = (\log\phi_i, \log c_i, \log\delta_i, \log\lambda_i, \log\rho_i, \log N_i, \log k_i)^T$, $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_i, i = 1, \cdots, n\}$, $\boldsymbol{\Theta}_{\{i\}} = \{\boldsymbol{\theta}_l, l \neq i\}$ and $\mathbf{Y} = \{y_{ij}, i = 1, \cdots, n;\ j = 1, \cdots, m_i\}$. Let $f_{ij}(\boldsymbol{\theta}_i, t_j) = \log_{10}(V_i(\boldsymbol{\theta}_i, t_j))$, where $V_i(\boldsymbol{\theta}_i, t_j)$

denotes the numerical solution of $V(t)$ in the differential equations (10) for the ith subject at time t_j. The repeated measurements of common logarithmic viral load for the ith subject, $y_{ij}(t)$, at treatment times t_j $(j = 1, 2, \cdots m_i)$ can be written as

$$y_{ij}(t_j) = f_{ij}(\boldsymbol{\theta}_i, t_j) + e_i(t_j), \;\; i = 1, \cdots n; \;\; j = 1, \cdots m_i \tag{12}$$

where $e_i(t)$ is a measurement error with mean zero. Note that log-transformation of dynamic parameters and viral load is used to make sure estimates of dynamic parameters to be positive and to stabilize the variance, respectively. Then the Bayesian nonlinear mixed-effects model can be written as the following three stages [12, 19, 63].

Stage 1. Within-subject variation:

$$\mathbf{y}_i = \mathbf{f}_i(\boldsymbol{\theta}_i) + \mathbf{e}_i, \qquad [\mathbf{e}_i|\sigma^2, \boldsymbol{\theta}_i] \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{m_i}) \tag{13}$$

where $\mathbf{y}_i = (y_{i1}(t_1), \cdots, y_{im_i}(t_{m_i}))^T$, $\mathbf{f}_i(\boldsymbol{\theta}_i) = (f_{i1}(\boldsymbol{\theta}_i, t_1), \cdots, f_{im_i}(\boldsymbol{\theta}_i, t_{m_i}))^T$, $\mathbf{e}_i = (e_i(t_1), \cdots, e_i(t_{m_i}))^T$ and the bracket notation $[A|B]$ denotes the conditional distribution of A given B.

Stage 2. Between-subject variation:

$$\boldsymbol{\theta}_i = \boldsymbol{\mu} + \mathbf{b}_i, \qquad [\mathbf{b}_i|\boldsymbol{\Sigma}] \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \tag{14}$$

Stage 3. Hyperprior distributions:

$$\sigma^{-2} \sim Ga(a, b), \quad \boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Lambda}), \quad \boldsymbol{\Sigma}^{-1} \sim Wi(\boldsymbol{\Omega}, \nu) \tag{15}$$

where the mutually independent Gamma (Ga), Normal ($\mathcal{N}$) and Wishart (Wi) prior distributions are chosen to facilitate computations [12, 19]. Note that the parametrization of the Gamma and Wishart distributions is such that $Ga(a, b)$ has mean ab and $Wi(\boldsymbol{\Omega}, \nu)$ has mean matrix $\nu\boldsymbol{\Omega}$. The hyper-parameters $a, b, \boldsymbol{\eta}, \boldsymbol{\Lambda}, \boldsymbol{\Omega}$ and ν are known.

Following the studies in [12, 19, 62, 63], it is shown from (13)–(15) that the full conditional distributions for the parameters $\sigma^{-2}, \boldsymbol{\mu}$ and $\boldsymbol{\Sigma}^{-1}$ can be written explicitly as

$$[\sigma^{-2}|\boldsymbol{\mu}, \boldsymbol{\Sigma}^{-1}, \boldsymbol{\Theta}, \mathbf{Y}] \sim Ga\left(a + \frac{\Sigma_{i=1}^n m_i}{2}, \boldsymbol{A}^{-1}\right) = \pi(\sigma^{-2}|\boldsymbol{\mu}, \boldsymbol{\Sigma}^{-1}, \boldsymbol{\Theta}, \mathbf{Y}) \tag{16}$$

$$[\boldsymbol{\mu}|\sigma^{-2}, \boldsymbol{\Sigma}^{-1}, \boldsymbol{\Theta}, \mathbf{Y}] \sim \mathcal{N}\left(\boldsymbol{B}^{-1}\boldsymbol{C}, \boldsymbol{B}^{-1}\right) = \pi(\boldsymbol{\mu}|\sigma^{-2}, \boldsymbol{\Sigma}^{-1}, \boldsymbol{\Theta}, \mathbf{Y}) \tag{17}$$

$$[\boldsymbol{\Sigma}^{-1}|\sigma^{-2}, \boldsymbol{\mu}, \boldsymbol{\Theta}, \mathbf{Y}] \sim Wi\left(\boldsymbol{D}^{-1}, n + \nu\right) = \pi(\boldsymbol{\Sigma}^{-1}|\sigma^{-2}, \boldsymbol{\mu}, \boldsymbol{\Theta}, \mathbf{Y}) \tag{18}$$

where $\boldsymbol{A} = b^{-1} + \frac{1}{2}\Sigma_{i=1}^n \Sigma_{j=1}^{m_i}[y_{ij} - f_{ij}(\boldsymbol{\theta}_i, t_j)]^2$, $\boldsymbol{B} = n\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda}^{-1}$, $\boldsymbol{C} = \boldsymbol{\Sigma}^{-1}\Sigma_{i=1}^n \boldsymbol{\theta}_i + \boldsymbol{\Lambda}^{-1}\boldsymbol{\eta}$, and $\boldsymbol{D} = \boldsymbol{\Omega}^{-1} + \Sigma_{i=1}^n(\boldsymbol{\theta}_i - \boldsymbol{\mu})(\boldsymbol{\theta}_i - \boldsymbol{\mu})^T$. The full conditional distribution of each $\boldsymbol{\theta}_i$, given the

remaining parameters and the data, can not be written explicitly, but it can be seen that the density function of the conditional distribution of $[\theta_i|\sigma^{-2}, \mu, \mathbf{\Sigma}^{-1}, \mathbf{\Theta}_{\{i\}}, \mathbf{Y}]$ is proportional to

$$\exp\left\{\frac{-\sigma^{-2}}{2}\Sigma_{j=1}^{m_i}[y_{ij} - f_{ij}(\boldsymbol{\theta}_i, t_j)]^2 - \frac{1}{2}(\boldsymbol{\theta}_i - \boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\boldsymbol{\theta}_i - \boldsymbol{\mu})\right\} = \pi(\boldsymbol{\theta}_i|\sigma^{-2}, \boldsymbol{\mu}, \mathbf{\Sigma}^{-1}, \mathbf{\Theta}_{\{i\}}, \mathbf{Y}) \tag{19}$$

4.2. MCMC Implementation

To carry out the Bayesian inference, it remains to specify the values of the hyper-parameters in the prior distributions (15). In the Bayesian approach, we only need to specify the priors at the population level which are easy to obtain from previous studies or literatures and usually are accurate and reliable. The population prior information also helps to resolve the identifiability problem at the individual level.

In principle, if we have reliable prior information for some of the parameters, strong priors with small variances may be used for these parameters. For other parameters such as ϕ, we may not have enough prior information or we may intend to use the available clinical data to determine since they are critical to quantify between-subject variations in response; a non-informative prior with large variance may be given for these parameters. In particular, one usually chooses non-informative prior distributions for parameters of interest [9].

After we specify the model for the observed data and the prior distributions for the unknown parameters, we can draw statistical inference for the unknown parameters based on their posterior distributions. In above Bayesian modeling approach, evaluation of the required posterior distributions in a closed-form is prohibitive. However, as indicated in Section 3.1, it is relatively straightforward to derive either full conditional distributions for some parameters or explicit expressions which are proportional to the corresponding full conditional distributions for other parameters.

Under the Bayesian framework, MCMC methods enable us to draw samples from the target distributions of interest or the posterior distributions of unknown parameters. In this paper we combine both the Gibbs sampler and the Metropolis-Hasting (M-H) algorithm to carry out the MCMC procedure. For more detailed discussion of this specific MCMC algorithms, please refer to literature [3, 18, 19, 22, 35, 57, 62, 63]. In our approach, the Gibbs sampling steps update $\sigma^{-2}, \boldsymbol{\mu}$ and $\mathbf{\Sigma}^{-1}$, while the M-H algorithm updates $\boldsymbol{\theta}_i,\ i = 1, \cdots, n$. After collecting the final MCMC samples, we are able to draw statistical inference for the unknown parameters. In particular, we are interested in the posterior means and quantiles.

To implement the M-H algorithm, it is necessary to specify a suitable proposal density. Several possible choices of proposal density are discussed in literatures, and a popular choice is the multivariate normal distribution, which results in the random-walk M-H algorithm [3, 9, 22, 52]. In our implementation, the proposal density is chosen to be a multivariate normal

distribution centered at the current value of $\boldsymbol{\theta}_i$, as it can be easily sampled and is symmetric [9, 22, 52, 63]. An important issue regarding the random-walk M-H algorithm is the choice of the dispersion of the proposal density. If the variance of the proposed density is too large, a large proportion of proposed moves will be rejected, and the Markov chain will therefore produce many repeats and result in inefficiency of the algorithm. On the other hand, if the variance of the proposed density is too small, the chain will have a high acceptance rate but will move around the parameter space slowly, again leading to inefficiency [9, 18, 52]. We will consider this issue in the MCMC implementation.

As suggested by Geman and Geman ([23]), for example, one long run may be more efficient with considerations of the following two points: (1) a number of initial "burn-in" simulations are discarded, since from an arbitrary starting point it would be unlikely that the initial simulations came from the stationary distribution targeted by the Markov chain; (2) one may only save every kth (k being an integer) simulation samples to reduce the dependence among samples used for parameter estimation. We are going to use these strategies in our MCMC implementation.

The iterative MCMC algorithm is outlined as follows.

Step 1. Initialize the iteration counter of the chain $j = 1$ and start with initial values $\Gamma^{(0)} = (\sigma^{-2(0)}, \boldsymbol{\mu}^{(0)}, \boldsymbol{\Sigma}^{-1(\mathbf{0})}, \boldsymbol{\Theta}^{(0)})^T$.

Step 2. Obtain a new value $\Gamma^{(j)} = (\sigma^{-2(j)}, \boldsymbol{\mu}^{(j)}, \boldsymbol{\Sigma}^{-1(j)}, \boldsymbol{\Theta}^{(j)})^T$ from $\Gamma^{(j-1)}$ through successive generation of values:

1. $\sigma^{-2(j)} \sim \pi(\sigma^{-2}|\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\Sigma}^{-1(j-1)}, \boldsymbol{\Theta}^{(j-1)}, \mathbf{Y})$
   $\boldsymbol{\mu}^{(j)} \sim \pi(\boldsymbol{\mu}|\sigma^{-2(j)}, \boldsymbol{\Sigma}^{-1(j-1)}, \boldsymbol{\Theta}^{(j-1)}, \mathbf{Y})$
   $\boldsymbol{\Sigma}^{-1(j)} \sim \pi(\boldsymbol{\Sigma}^{-1}|\sigma^{-2(j)}, \boldsymbol{\mu}^{(j)}, \boldsymbol{\Theta}^{(j-1)}, \mathbf{Y})$
2. For $\boldsymbol{\theta}_i^{(j)}$, move the chain to a new value φ, generated from the proposal density $q(\varphi|\boldsymbol{\theta}_i^{(j-1)})$, from $\boldsymbol{\theta}_i^{(j-1)}$. Evaluate the acceptance probability of the move, $\alpha(\varphi|\boldsymbol{\theta}_i^{(j-1)})$ given by (20). If the move is accepted, $\boldsymbol{\theta}_i^{(j)} = \varphi$. If it is not accepted, $\boldsymbol{\theta}_i^{(j)} = \boldsymbol{\theta}_i^{(j-1)}$ and the chain does not move.

Step 3. Change the counter from j to $j + 1$ and return to Step 2 until convergence is reached.

In Step 2.2 a sample u is drawn from the uniform distribution with support on $(0, 1)$. If $u \leq \alpha(\varphi|\boldsymbol{\theta}_i^{(j-1)})$, the proposed move is accepted and if $u > \alpha(\varphi|\boldsymbol{\theta}_i^{(j-1)})$, the move is not allowed, where

$$\alpha(\varphi|\boldsymbol{\theta}_i^{(j-1)}) = min\left\{1, \frac{\pi(\varphi|\sigma^{-2(j)}, \boldsymbol{\mu}^{(j)}, \boldsymbol{\Sigma}^{-1(j)}, \boldsymbol{\Theta}_{\{i\}}^{(j-1)}, \mathbf{Y})}{\pi(\boldsymbol{\theta}_i^{(j-1)}|\sigma^{-2(j)}, \boldsymbol{\mu}^{(j)}, \boldsymbol{\Sigma}^{-1(j)}, \boldsymbol{\Theta}_{\{i\}}^{(j-1)}, \mathbf{Y})}\right\} \tag{20}$$

Note that the choice of the proposal density q is essentially arbitrary, although in practice a careful choice will help the algorithm to move quickly around the parameter space. Due to the

symmetry of q, $q(\varphi|\boldsymbol{\theta}_i^{(j-1)}) = q(\boldsymbol{\theta}_i^{(j-1)}|\varphi)$ and they are cancelled from fractional expression in (20). For $\boldsymbol{\theta}_i$, we use a multivariate normal proposal distribution centered at the current value with variance-covariance matrix given by an information-type matrix [3, 22, 24, 63]. In addition, for each iteration, we need to derive $f_{ij}(\boldsymbol{\theta}_i, t_j)$ by solving the differential equations (10) so that samples of σ^{-2} and $\boldsymbol{\theta}_i$ can be drawn from (16) and (19), respectively.

In the implementation of simulation studies and the actual clinical data application below, an informal check of convergence is conducted based on graphical techniques according to the suggestion of Gelfand and Smith [20]. An example of graphical results will be displayed in Section 4 below. Based on the results, we propose that, after an initial number of 30,000 burn-in iterations, every 5th MCMC sample was retained from the next 120,000 samples. Thus, we obtained 24,000 samples of targeted posterior distributions of the unknown parameters.

4.3. Sensitivity Analysis

As we have known, a common concern with Bayesian methods is their dependence on various aspects of the modeling process. Possible sources of uncertainty include the prior distributions, the number of levels in the hierarchical model and the initial values. The basic tool for investigating model uncertainty is the sensitivity analysis. That is, we simply make reasonable modifications to the assumptions in question, recompute the posterior quantities of interest, and see whether they have changed in a way that significantly affects the resulting interpretations or conclusions. If the results are robust against the suspected assumptions, we can report the results with confidence and our conclusions will be solid. However, if the results are sensitive to the assumptions, we wish to communicate the sensitivity results and interpret the results with caution [9]. We will investigate the sensitivity of our models in the simulation studies (example 1 below).

5. SIMULATION STUDIES

In this section, we present two simulated numerical examples to illustrate the introduced Bayesian approach. The scenario we consider is as follows. For each example, we simulate a clinical trial with 20 HIV infected patients in a long-term period of antiretroviral treatment. For each patient, we assume that measurements of viral load are taken at days 0, 1, 2, 4, 6, 8, 12, 14, 16, 20, 26, 32, 38, 44, 56, 64, 72, 80, 90, 100, 120, 140, 160, 180 and 200 of follow-up. The design of this experiment is similar to an actual AIDS clinical trial that we will describe in details latter, but here more frequent data are simulated for the purpose of easy implementation.

Because the parameters of the model (13) are not identifiable, meaning that different combinations of parameters lead to the same likelihood, it is impossible to obtain unique maximum likelihood estimates of parameters based on the data. A potential advantage of Bayesian analysis over likelihood method is that if an informative prior is available, then Bayesian inferences

can be obtained despite the fact that a model is not identifiable (from the perspective of the likelihood) [51]. Based on this consideration, we design the following two examples to illustrate our approach in order to handle the problems of parameter identifiability. In the first example, it can be seen that the two parameter $\log c$ and $\log \delta$ can be identified if we assume that the other five parameters $(\log \phi, \log \lambda, \log \rho, \log N, \log k)$ are constant. Since the classic methods of identifiability [2, 11, 29, 34, 59] for a system of nonlinear differential equations can not be used, and the exact identifiability checks for nonlinear differential equation models are unfortunately not available [2], an informal check of parameter identifiability based on graphical techniques can be used by studying samples drawn from the MCMC sampling scheme for each parameter. We can check the k-lag serial correlation of the samples for each parameter [7, 14]. If the model is unidentifiable and the prior distribution is not informative, the k-lag serial correlation tends to be large even for a large k. In this case, the trace plot usually lacks randomness, i.e., the consecutive samples move towards one direction. In practice, one can use these facts to check the identifiability of parameters by carefully studying the samples drawn from the MCMC scheme [21]. Sampling-based series correlation check can not only detect the possible unidentifiability, but also shed some light on the relationship among unidentifiable parameters. In example 2, we consider all the seven parameters to be unknown, but we assume that $\log \phi$ has a non-informative prior and the other six parameters have informative priors.

Based on the discussion in Section 4.2, the prior distribution for $\boldsymbol{\mu}$ was assumed to be $\mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Lambda})$ with $\boldsymbol{\Lambda}$ being diagonal matrix. Following the idea of Han *et al.* [24] for prior construction, as an example we discuss the prior construction for $\log \delta$. The prior constructions for other parameters are similar and so are omitted here.

Ho *et al.* [25] reported viral dynamic data on 20 patients; the logarithm of the average death rate of infected cells ($\log \delta$) was -1.125. Wei *et al.* [64] used two different models with a group of 22 subjects to estimate death rate of infected cells and obtained $\log \delta$ with –0.84 and –1.33, respectively. Following the studies by [25, 64], Nowak *et al.* [40] estimated $\log \delta = -0.934$ based on 11 subjects with one possible outlying subject excluded. The individual estimates of $\log \delta$ from these studies approximately follow a symmetric normal distribution. Thus we chose a normal distribution $\mathcal{N}$(–1.0, 0.0025) as the prior for $\log \delta$ (the small variance indicated that we used an informative prior for $\log \delta$).

In this study, the values of the hyper-parameters were determined based on several studies in literature [24, 25, 42, 43, 44, 45, 46, 58, 64]. In addition, the data for the pharmacokinetic factor (C_{12h}), phenotype marker (baseline and failure $IC_{50}s$) and adherence as well as the baseline viral load (V_0) were taken from an AIDS clinical trial study (Section 6).

5.1. Example 1

In this example, we design an experiment to only estimate the two parameters $\log c$ and $\log \delta$

which are identifiable in our model, and assume that the other five parameters are given as constants whose values are $(\log\phi, \log\lambda, \log\rho, \log N, \log k) = (2.5,\ 4.6,\ -2.3,\ 6.9,\ -11.0)$. Based on the discussion in Section 4.2, the prior distribution for $\boldsymbol{\mu} = (\log c, \log\delta)^T$ was assumed to be $\mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Lambda})$ with $\boldsymbol{\Lambda}$ being diagonal matrix. Thus, the values of hyper-parameters are chosen as follows

$a = 4.5,\ \ b = 9.0,\ \ \nu = 5.0,\ \ \boldsymbol{\eta} = (1.1,\ -1.0)^T,\ \ \boldsymbol{\Lambda} = \mathrm{diag}(1000.0,\ 1000.0),\ \ \boldsymbol{\Omega} = \mathrm{diag}(2.5,\ 2.5).$

Note that the non-informative priors are chosen for both $\log c$ and $\log\delta$.

The true individual dynamic parameters, $\log c_i$ and $\log\delta_i$, are generated by $\log c_i = \log c + b_{1i}$ and $\log\delta_i = \log\delta + b_{2i}$, where $\log c = 1.1$ and $\log\delta = -1.0$ are the true values of population parameters, both b_{1i} and b_{2i} are random effects following a normal distribution with mean 0 and standard deviation of 0.2.

Based on generated true parameters, known parameters and data (C_{12h}, IC_{50} and $A(t)$), the observations y_{ij} (the common logarithm of total viral load) are generated by perturbing the solution of the differential equations (10) with a within-subject measurement error, *i.e.*, $y_{ij} = \log_{10}(V_{ij}) + e_i$, where V_{ij} is the numerical solution of viral load to the differential equations (10) for the ith subject at time t_j. It is assumed that the within-subject measurement error e_i is normally distributed with $\mathcal{N}(0, 0.15^2)$. A more complicated model with random measurement errors may also be used. For example, we may use the multiplicative error or model the variance of the error as a function of the mean [12].

We apply the introduced Bayesian approach to estimate the dynamic parameters. As discussed previously, the graphical check of identifiability in parameters based on the last 500 samples drawn from MCMC sampling scheme for both parameters $\log c$ and $\log\delta$ is presented in Figure 2. It can be seen that the consecutive samples move randomly towards different directions which indicates that MCMC sampler is not "sticky" and the two parameters are regarded as to be identifiable.

Place Figure 2 here

Figure 3 displays the three representative individual fitted curves with generated viral load data in $\log_{10}$ scale, the estimated drug efficacy ($\hat{\gamma}(t)$) with threshold (e_c), as well as observed $IC_{50}(t)$ and adherence of the two PI drugs. It can be seen that the models provide a good fit to the generated data. It is worth noting, by comparing the plots of fitted curves and estimated drug efficacy $\hat{\gamma}(t)$, that if $\hat{\gamma}(t)$ falls below the threshold e_c, viral load rebounds, and in contrast, if $\hat{\gamma}(t)$ is above e_c, the corresponding viral load does not rebound, which is consistent with our theoretical analysis of the dynamic models [27].

Place Figure 3 here

In Table 1, we summarize the generated true values of parameters ($\log c$ and $\log \delta$) and the mean estimates with 40 replications for the 20 subjects, as well as the corresponding bias, which is the difference between mean estimate and true value of parameters, and standard error (SE), defined as the square root of mean-squared error. The percentage is based on the absolute value of the true parameter. It can be seen from Table 1 that both bias and the SE for population parameter estimates are very small. For individual parameter estimates, the bias is also small, ranging from 0.001 to 0.243, and the SE (%) ranges from 0.6 to 25.7.

Place Table 1 here

In order to examine the dependence of dynamic parameter estimates on the prior distributions and initial values, we carried out the sensitivity analysis. We follow the method proposed by Raftery and Lewis [50] to implement the MCMC sampling scheme using FORTRAN codes and monitor several independent MCMC runs, starting from different initial values. Those runs exhibit similar and stable behavior. An informal check of convergence diagnostics based on graphical techniques suggested in [20] is investigated. As an example, the number of MCMC iterations and convergence with regard to three different initial values are displayed in Figure 4. In addition, compared with the mean vector $\boldsymbol{\eta} = (1.1,\ -1.0)^T$ of prior distributions which was used to obtain the results shown in Figure 3 and Table 1, we chose two alternative mean vectors $\boldsymbol{\eta} = (1.3,\ -0.8)^T$ (higher level) and $\boldsymbol{\eta} = (0.9,\ -1.2)^T$ (lower level) of prior distributions, respectively, with three sets of different initial values in the sensitivity analysis (data not shown). We summarize the sensitivity analysis results as follows: (i) The estimated dynamic parameters were not sensitive to both the priors and/or the initial values, and results are reasonable and robust. (ii) When different priors and/or different initials were used, the results are similar to those presented in Figure 3 and Table 1.

Place Figure 4 here

5.2. Example 2

We design this experiment to consider all the seven parameters to be estimated. Similar to example 1, the prior distribution for $\boldsymbol{\mu} = (\log \phi, \log c, \log \delta, \log \lambda, \log \rho, \log N, \log k)^T$ is assumed to be $\mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Lambda})$ with $\boldsymbol{\Lambda}$ being diagonal matrix. We chose the values of the hyper-parameters as

follows:

$$a = 4.5,\ \ b = 9.0,\ \ \nu = 8.0,\ \ \boldsymbol{\eta} = (2.5,\ 1.1,\ -1.0,\ 4.6,\ -2.5,\ 6.9,\ -11.0)^T,$$
$$\boldsymbol{\Lambda} = \text{diag}(1000.0,\ 0.0025,\ 0.0025,\ 0.0025,\ 0.0025,\ 0.0025,\ 0.001),$$
$$\boldsymbol{\Omega} = \text{diag}(1.25,\ 2.5,\ 2.5,\ 2.0,\ 2.0,\ 2.0,\ 2.0).$$

We generate the true parameters using the between-subject variation model (14), $\boldsymbol{\theta}_i = \boldsymbol{\mu} + \mathbf{b}_i$ $(i = 1, \cdots, 20)$, where we assume that the population parameter vector $\boldsymbol{\mu} = (2.5,\ 1.1,\ -1.0,\ 4.6,\ -2.5,\ 6.9,\ -11.0)^T$ and the random effects $\mathbf{b}_i$ are normally distributed with mean $\mathbf{0}$ and diagonal standard deviation matrix of $diag(0.1,\ 0.2,\ 0.2,\ 0.1\ 0.2,\ 0.1,\ 0.1)$. In the same way as we did in Example 1, we generate the observations y_{ij} (the common logarithm of viral load) to estimate the viral dynamic parameters by applying the Bayesian approach introduced in Section 4. The experiment is replicated 40 times. Similar to Example 1, the model provides a good fit to the generated data (plots not shown here).

Table 2 summarizes the bias and standard error (SE) for population (Pop) and individual dynamic parameters with 40 replications. Again, the percentage is based on the absolute value of the true parameter. It can be seen from the results that the bias of estimates is very small (for the population parameter estimates, the bias ranged from 0.001 to 0.023; for the individual estimates, the bias ranged from 0.001 to 0.355). The standard error is reasonable (ranged from 1.5% to 15.3% for population estimates and from 0.8% to 39.7% for individual estimates).

Place Table 2 here

6. APPLICATION TO AN AIDS CLINICAL TRIAL STUDY

We apply the proposed methodology to the data from an AIDS clinical study. This study was a Phase I/II, randomized, open-label, 24-week comparative study of the pharmacokinetic, tolerability and antiretroviral effects of two regimens of indinavir (IDV), ritonavir (RTV), plus two nucleoside analogue reverse transcriptase inhibitors (NRTIs) on HIV-1-infected subjects failing PI-containing antiretroviral therapies [1]. The 44 subjects were randomly assigned to the two treatment Arm A (IDV 800 mg q12h + RTV 200 mg q12h) and Arm B (IDV 400 mg q12h + RTV 400 mg q12h). Out of the 44 subjects, 42 subjects are included in the analysis; for the remaining two subjects, one was excluded from the analysis since the PK parameters were not obtained and the other was excluded since PhenoSense HIV could not be completed on this subject due to an atypical genetic sequence that causes the viral genome to be cut by an enzyme used in the assay. Plasma HIV-1 RNA (viral load) measurements were taken at days 0, 7, 14, 28, 56, 84, 112, 140 and 168 of follow-up. The data for pharmacokinetic parameters (C_{12h}), phenotype marker (baseline and failure $IC_{50}s$) and adherence from this study were also

used in our modeling. The adherence data were determined according to pill-count data. More detailed description of this study can be found in publication by Acosta *et al.* [1].

To implement the Bayesian approach, we used the same values of the hyper-parameters as those in Example 2. The MCMC techniques consisting of a series of Gibbs sampling and M-H algorithms were used to obtain the results presented as follows.

Figures 5 shows three individually fitted curves with observed viral load data in $\log_{10}$ scale, the corresponding estimated drug efficacy function ($\hat{\gamma}(t)$) with threshold (e_c), the observed $IC_{50}(t)$ and the adherence, $A(t)$, of the two PI drugs. It can be seen that the model provides a good fit to the observed data for these three subjects. It is also found by comparing the plots of fitted curves and estimated drug efficacies that, in general, if $\hat{\gamma}(t)$ falls below the threshold e_c, viral load rebounds, and in contrast, if $\hat{\gamma}(t)$ is above e_c, the corresponding viral load does not rebound. This further confirms our theoretical results for the viral dynamic models [27].

Place Figure 5 here

A large between-subject variation in the estimates of all individual dynamic parameters is observed (data not shown here). The population posterior means and the corresponding 95% equal-tail credible intervals for the seven parameters are summarized in Table 3. It is shown that the population estimates are 3.06 and 0.37 for c and δ, respectively, which are the most important parameters in understanding viral dynamics. In comparison with previous studies, our population estimate of c (3.06) is almost equal to the mean estimate of c, 3.07 in [44], and our population estimate of δ is consistent with the mean value of δ, 0.37 in [33, 58]. However, our population estimate of c is slightly less than the mean estimate of c, 3.1 in [46] and is greater than the population estimate of c, 2.81, with credible interval being (1.24, 6.49) obtained by Han *et al.* [24]. On the other hand, our population estimate of δ (0.37) is less than the first-phase decay rate of 0.43 [40], 0.49 [44] and 0.5 [46]. In addition, in two separate studies by Markowitz *et al.*[36] and Perelson *et al.*[45], the mean values of 1.0 and 0.7 for δ were obtained by holding clearance rate c as constant with values of 23 and 3, respectively, and these two values are substantially greater than our population estimate of 0.37 for δ. These differences may be due to various reasons as follows. The analysis of those studies assumed that viral replication was completely stopped by the treatment, and/or they used short-term viral load data to fit their models. In addition, the first-phase decay rate, estimated from a biexponential viral dynamic model [25, 44, 66] under perfect treatment assumption, is not the exact death rate of infected cells (δ) since the current antiretroviral therapy cannot completely block viral replication [8, 13, 46]. In this study, we estimated the death rate of infected cells (δ) directly by accounting for the non-perfect treatment with time-varying drug efficacy. Note that we are

unable to validate our results of the other parameter estimates as no conclusive or comparable estimates have been published to date.

Place Table 3 here

It is also important that we can estimate the threshold of the drug efficacy (e_c). The efficacy threshold may represent how good the immune status of a patient can control viral replication. If the efficacy threshold (e_c) for a patient is small, it may indicate that this patient's immune response to the virus is strong and a regimen with a mild potency can keep the virus on check for this patient. However, if the efficacy threshold (e_c) for a patient is high, it indicates that this patient needs a highly potent regimen to suppress the virus. Thus, the efficacy threshold (e_c) is important for individual patients.

7. DISCUSSION

In this paper, we propose a concept of longitudinal dynamic systems, in particular, for modeling HIV dynamics. Our models are simplified with the main goals to retain crucial features of HIV-1 dynamics and, at the same time, to guarantee their applicability to typical clinical data, in particular, total viral load measurements. We investigated a hierarchical Bayesian (mixed-effects) modeling approach to estimate dynamic parameters in the proposed mathematical model for long-term HIV dynamics. Fitting of mathematical models using a Bayesian approach is a powerful way in analyzing data from studies of viral dynamics. First, Bayesian modeling involves specifying prior distributions of model parameters to perform the analysis. Thus, it can not only incorporate the estimates of dynamic parameters from previous studies, but also handle the identifiability problems of parameters. Secondly, the Bayesian approach allows the fitting of complex models for viral dynamics and is more flexible than other methods such as nonlinear least squares (NLS) method. Thirdly, the graphical output of simulation-based Bayesian algorithms provides both informative diagnostic aids and easily understood inferential summaries.

We have presented two simulation examples and an actual AIDS clinical trial study to illustrate how the Bayesian procedures can be applied to HIV dynamic studies. Both the population and individual dynamic parameters can be estimated from the hierarchical Bayesian modeling approach. For simulation studies, it was seen that the models provided a good fit to the simulated data. The bias for both population and individual dynamic parameters is very small, and the SE (%) of the estimates is reasonable. One thus might claim that both population and individual parameters would be identifiable by only providing the population prior information under a framework of the hierarchical Bayesian model based on our simulation studies. For the

actual AIDS clinical trial data set, the proposed model fitted the clinical data reasonably well for most subjects in our study, although the fitting for a few subjects (less than 10%) was not completely satisfactory due to unusual viral load response patterns, inaccurate measurements of drug exposure and/or adherence for these subjects. For example, the measurement of adherence may not reflect actual adherence profiles for the individual subject.

Most of previous studies have assumed perfect drug effect [25, 36, 40, 44, 45, 46, 65] or imperfect constant drug effect [15, 46, 65] to estimate dynamic parameters with short-term viral load data. These assumptions contributed to the limitations of those studies which might result in inaccuracy of dynamic parameter estimation. Compared with those studies, our model proposed in this paper has the following features: (i) time-varying drug efficacy during long-term treatment; (ii) more reasonable biological interpretation; (iii) incorporating drug concentration, adherence and resistance in the model; and (iv) good fit to the observed long-term viral load data (whole data). Thus, based on this model, the results of estimated dynamic parameters should be more reliable and reasonable to interpret long-term HIV dynamics.

As indicated previously, it was interesting to note that, by comparing the results of fitted curves and estimated drug efficacy $\hat{\gamma}(t)$ (obtained in combination with the clinical data C_{12h}, $A(t)$, $IC_{50}(t)$ and the estimated values of ϕ), if $\hat{\gamma}(t)$ falls below the threshold e_c, viral load rebounds, and in contrast, if $\hat{\gamma}(t)$ is above e_c, the corresponding viral load does not rebound. Thus, the threshold of drug efficacy e_c may reflect ability of the immune system of a patient for controlling virus replications. It is therefore important to estimate e_c for each patient based on clinical data.

Although the analysis presented here used a simplified model which appeared to perform well in capturing and explaining the observed patterns, and characterizing the biological mechanisms of HIV infection under relatively complex clinical situations, our model is however limited in several ways. Our mathematical model (10) is a simplified model among many variations of viral dynamic models [8, 42, 46]. We did not consider the compartments of productively infected cells, long-lived and latently infected cells separately [45]. Instead we pooled all the infected cell populations together. The virus compartment was not further decomposed into infectious virions and non-infectious virions as in Perelson *et al.* [44]. Thus, different mechanisms of NRTI and PI drug effects were not modeled. In fact, we only considered the PI drug effects in the drug efficacy model (9) since the information of NRTI drugs was not collected in our study and the effect of NRTI drugs was considered less important compared to the PI drugs. We modeled the drug resistance using the phenotype IC_{50} values instead of modeling genotype viral species separately [42]. One of the main reasons is that genotypic assays is hard to interpret due to the large number of mutations that lead to resistance of antiretroviral drugs. Although more elaborated models with consideration of more infected cell and virus compartments, more detailed drug effects, and specific drug resistant viral species may provide more accurate descriptions for the long-

term HIV dynamics, they may give rise to the identifiability problems of model parameters due to the complexity of the models, and thus limit the usefulness of these models. The trade-off between the complexity and applicability of HIV dynamic models should be considered, and further studies on this issue are definitely needed. Nevertheless, these limitations would not offset the major findings from our modeling approach.

We assumed that the distribution of the random effects in (14) is normal. However, due to the nature of AIDS clinical data, it is possible that data may contain outlying individuals and, thus, may result in a skewed distribution of individual parameters, *i.e.* the random effects may not follow a normal distribution. As Wakefield [63] suggested, a t distribution may be used which is more robust to outlying individuals than the normal distribution. We plan to address this issue and report results in future studies.

In summary, the mechanism-based dynamic model is powerful and efficient to establish the relationship between antiviral response and drug exposure and drug susceptibility, although some biological assumptions have to be made. The fitting of a model specified as a set of differential equations to data is routinely done in many fields (in particular, pharmacokinetics and pharmacodynamics which are closely associated with the analysis of clinical data considered in this paper). Our hope is that this work might stimulate the investigation of more realistic models to analyze data from AIDS clinical trials with antiviral treatment which, in turn, would help to better understand the biological mechanisms of HIV infection, to study the pathogenesis of AIDS progression, to guide development of antiviral treatment strategies and to take into account the roles of clinical factors in antiviral activities. We also expect that the proposed concept of longitudinal dynamic systems can be applied to other biological processes.

Acknowledgments

We thank Drs. John G. Gerber, Edward P. Acosta and other A5055 study investigators for their collaborations and allowing us to use the clinical data from their study. The authors are also indebted to Dr. Alan S. Perelson from Los Alamos National Laboratory and Professor Jun S. Liu from Harvard University for their informative discussions. This work was supported in part by National Institutes of Health (NIH) research grants RO1 AI052765 and RO1 AI055290.

References

[1] Acosta, E., P., Wu, H., Hammer, S. M., *et al.* Comparison of two indinavir/ritonavir regimens in the treatment of HIV-infected individuals. *JAIDS*, in press.

[2] Audoly, S. *et al.* (2001), "Global identifiability of non-linear models of biological systems," *IEEE Trans. Biomed. Eng.*, **48**, 55–65.

[3] Bennett, J. E., Racine-Poon, A., and Wakefield, J. C. (1996). Markov chain Monte Carlo for nonlinear hierarchical models, in *Markov chain Monte Carlo in Practice*. Gilks WR, Richardson, S., Spiegelhalter, D.J. (Eds), Chapman & Hall, London, 339–357.

[4] Besag, J., Green, P. J. (1993). Spatial statistics and Bayesian computation. *Journal of the Royal Statistical Society B* **55**, 25–37.

[5] Besch, C. L. (1995). Compliance in clinical trials. *AIDS*, **9**, 1–10.

[6] Boeckmann, A. J., Beal, S. L., Sheiner, L. B. (1989). *NONMEM Users Guides*. Technical report, Division of Clinical Pharmacology, University of California at San Francisco.

[7] Brockwell, P. J., Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd. Springer-Verlag: New York.

[8] Callaway, D. S., and Perelson, A. S., (2002). HIV-1 infection and low steady state viral loads. *Bulletin of mathematical Biology* **64**, 29–64.

[9] Carlin, B. P., and Louis, T. A. (1996). *Bayes and empirical Bayes methods for data analysis*. London: Chapman & Hall.

[10] Chen,T., He, H. L., Church, G. M. (1999). Modeling gene expression with differential equations. *Proc. of Pacific Symposium of Biocomputing* **4**, 29–40.

[11] Cobelli, C., Lepschy. A., Jacur, G. R. (1979). Identifiability of compartmental systems and related structural properties. *Mathematical Biosciences* **44**, 1–18.

[12] Davidian, M., and Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. London: Chapman & Hall.

[13] Ding, A. A., and Wu, H. (1999). Relationships between antiviral treatment effects and biphasic viral decay rates in modeling HIV Ddynamics. *Mathematical Biosciences* **160**, 63–82.

[14] Diggle, P. J. (1990). *Time Series: A Biostatistical Introduction*. Oxford: Oxford University Press.

[15] Ding, A. A., and Wu, H. (2000). A comparison study of models and fitting procedures for biphasic viral decay rates in viral dynamic models. *Biometrics* **56**, 16–23.

[16] Ding, A. A., and Wu, H. (2001). Assessing antiviral potency of anti-HIV therapies *in vivo* by comparing viral decay rates in viral dynamic models. *Biostatistics* **2**, 13–29.

[17] Galecki, A,T. (1998). NLMEM: a NEW SAS/IML macro for hierarchical nonlinear models. *Computer Methods and Programs in Biomedicine* **55**, 207–216.

[18] Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Londan: Chapman & Hall.

[19] Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F. M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association* **85**, 972–985.

[20] Gelfand, A. E., and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409.

[21] Gelfand, A. E., Sahu, S. K. (1999). Identifiability, improper priors, and Gibbs sampling for generalized linear models. *Journal of the American Statistical Association* **94**, 247–253.

[22] Gelman, A., Bois, F., and Jiang, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* **91**, 1400–1412.

[23] Geman, S., Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern Recognition and machine Intelligence* **6**, 721–741.

[24] Han, C., Chaloner, K., and Perelson, A. S. (2002). Bayesian analysis of a population HIV dynamic model. *Case Studies in Bayesian Statistics, Vol. 6*, New York: Springer-Verlag.

[25] Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., and Markowitz, M. (1996). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123–126.

[26] Hsu, A., Isaacson, J., Kempf, D. J. *et al.* (2000). Trough concentrations-EC_{50} relationship as a predictor of viral response for ABT-378/ritonavir in treatment-experienced patients. *40th Interscience Conference on Antimicrobial Agents and Chemotherapy*. San Francisco, CA, Poster session 171.

[27] Huang, Y., Rosenkranz, S. L., and Wu, H. (2003). Modeling HIV dynamics and antiviral responses with consideration of time-varying drug exposures, sensitivities and adherence. *Mathematical Biosciences* **184**, 165–186.

[28] Ickovics, J. R., and Meisler, A. W. (1997). Adherence in AIDS clinical trial: a framework for clinical research and clinical care. *Journal of Clinical Epidemiology* **50**, 385–391.

[29] Jacquez, J. J. and Greif, P. (1985). Numerical parameter identifiability and estimability: intergrating identifiability, estimability and optimal sampling design. *Mathematical Biosciences* **77**, 201–210.

[30] Jansson, B., Revesz, L. (1975). Analysis of the growth of tumor cell pupulations. *Mathematical Biosciences* **19**, 131–154.

[31] Kaufmann, G. R., *et al.* (1998). Patterns of viral dynamics during primary human immunodeficiency virus type 1 infection. *Journal of Infectiuos Diseases* **178**, 1812-1815.

[32] Kempf, D. J., Hsu, A., Jiang, P. *et al.* (2001). Response to ritonavir intensification in indinavir recipients is highly correlated with virtual inhibitory quotient. *8th Conference on Retroviruses and Opportunistic Infections*, Chicago, IL, Abstract 523.

[33] Klenernam, P., Phillips, R. E., *et al.* (1996). Cytotoxic T lymphocytes and viral turnover in HIV type 1 infrction. *Proc. Natl. Acad. Sci. USA* **93**, 15323–15328.

[34] Ljung, L., Glad, S. T., (1994). On global identifiability for arbitrary model parameterizations. *Automatica* **30**, 265–276.

[35] Lunn, D. J., Best, N., Thomas, A., Wakefield, J., and Spiegelhalter, D. (2002). Bayesian analysis of population PK/PD models: general concepts and software. *Journal of Pharmacokinetics and Pharmacodynamics* **29**, 271–307.

[36] Markowitz, M., Louie, M. *et al.* (2003). A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. *Journal of Virology* **77**, 5037-5038.

[37] Michelson, S., Leith, J. T. (1997). Tumor Heterogeneity and Growth Control, in *A survey of models for tumor-immune system dynamics*. Adam, J. A., Bellomo, N. (Eds), Birkhäuser, Boston, 295–326. 5037–5038.

[38] Molla, A. *et al.* (1996). Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nature Medicine* **2**, 760–766.

[39] Nelson, P. W., and Perelson, A. S. (2002). Mathematical analysis of delay differential equation models of HIV-1 infection. *Mathematical Biosciences* **179**, 73–94.

[40] Nowak, M. A., Bonhoeffer, S. *et al.* (1995). HIV results in the frame. *Nature* **375**, 193.

[41] Nowak, M. A., Bonhoeffer, S., Shaw, G. M., and May, R. M. (1997). Anti-viral drug treatment: dynamics of resistance in free virus and infected cell populations. *Journal of Theoretical Biology* **184**, 203–217.

[42] Nowak, M. A., and May, R. M. (2000). *Virus dynamics: mathematical principles of immunology and virology*. Oxford: Oxford University Press.

[43] Perelson, A. S., Kirschener, D. E., and Boer, R. D. (1993). Dynamics of HIV infection of $CD4^+$ T cells. *Mathematical Biosciences* **114**, 81–125.

[44] Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M., and Ho, D. D. (1996). HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582–1586.

[45] Perelson, A. S., Essunger, P. *et al.* (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* **387**, 188–191.

[46] Perelson, A. S., and Nelson, P. W. (1999). Mathematical analysis of HIV-1 dynamics *in vivo*. *SIAM Review* **41**(1), 3–44.

[47] Pinheiro, J., and Bates, D. M. (2000). *Mixed-effects models in S and S-plus*. New York: Springer

[48] Putter, H., Heisterkamp, S. H., Lange, J. M. A. and De Wolf, F.(2002). A Bayesian approach to parameter estimation in HIV dynamical models. *Statistics in Medicine* **21**, 2199–2214.

[49] Raftery, A. E., and Banfield, J. D. (1991). Stopping the Gibbs sampler, the use of morphology and other issues in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43**, 32–43.

[50] Raftery, A. E., Lewis, S. (1992). How many iterations in the Gibbs sample? in *Bayesian Statistics 4*. Bernardo J, Berger J, Dawid A, Smith A (Eds); Oxford: Oxford University Press, 763–773.

[51] Rannala, B. (2002). Identifiability of parameters in MCMC Bayesian inference of phylogeny. *Systematic Biology* **51**(5), 754–760.

[52] Roberts, G. O. (1996),. Markov chain concepts related to sampling algorithms. in *Markov chain Monte Carlo in Practice*. Gilks WR, Richardson S, Spiegelhalter DJ (Eds); London: Chapman & Hall, 45–57.

[53] SAS Institute Inc. (2000). *SAS/STAT User's Guide, Version 8*. SAS Publishing.

[54] Sheiner, L. B., Rosenberg, B., Marathe, K. L. (1977). Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *Journal of Pharmacokintics and Biopharmaceutics*. **5**, 635–651.

[55] Sheiner, L. B., and Beal, S. L. (1980). Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: Routine clinical pharmacokinetic data.*Journal of Pharmacokintics and Biopharmaceutics* **8**, 553–571.

[56] Sheiner, L. B. (1985). Modeling pharmacodynamics: parametric and nonparametric approaches, in *Variability in Drug Therapy: Description, Estimation, and Control*. Rowland M, *et al.* (Eds); New York: Raven Press, 139–152.

[57] Smith, A. F. M., and Roberts, G. O. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society, Series B* **55**, 3–23.

[58] Stafford, M. A. *et al.* (2000). Modeling plasma virus concentration during primary HIV infection. *Journal of Theoretical Biology* **203**, 285–301.

[59] Tunali, E. T., Tarn, T. J., (1987). New results for identifiability of nonlinear systems. *IEEE Trans. Automat. Contr.* AC-32, 146–154.

[60] Wahl, L. M., and Nowak, M. A. (2000). Adherence and resistance: predictions for therapy outcome. *Proceedings of the Royal Society, Biological* **267**, 835–843.

[61] Wainberg, M. A. *et al.* (1996). Effectiveness of 3TC in HIV clinical trials may be due in part to the M184V substition in 3TC-resistant HIV-1 reverse transcriptase. *AIDS*, 10(suppl), S3–S10.

[62] Wakefield, J. C., Smith, A. F. M., Racine-Poon, A., and Gelfand, A. E. (1994). Bayesian analysis of linear and non-linear population models using the Gibbs sampler. *Applied Statistics* **43**, 201–221.

[63] Wakefield, J. C. (1996). The Bayesian analysis to population Pharmacokinetic models. *Journal of the American Statistical Association* **91**, 62–75.

[64] Wei, X., Ghosh, S. K. *et al.* (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117–122.

[65] Wu, H., Ding, A. A., and de Gruttola, V. (1998). Estimation of HIV dynamic parameters. *Statistics in Medicine* **17**, 2463–2485.

[66] Wu, H., and Ding, A. A. (1999). Population HIV-1 dynamics *in vivo*: applicable models and inferential tools for virological data from AIDS clinical trials. *Biometrics* **55**, 410–418.

Table 1: The true values and mean estimates of population (Pop) and individual dynamic parameters with 40 replications as well as the corresponding bias and standard error (SE), defined as the square root of mean-squared error. The percentage of SE is based on the absolute value of the true parameter.

	True value		Mean estimate		Bias		SE (%)	
	$\log c$	$\log \delta$	$\log c$	$\log \delta$	$\log c$	$\log \delta$	$\log c$	$\log \delta$
Pop	1.100	-1.000	1.107	-1.006	0.007	-0.006	3.02	3.00
Sub1	0.658	-0.778	0.684	-0.767	0.026	0.011	17.1	21.6
Sub2	0.851	-1.216	0.865	-1.125	0.014	0.091	22.4	22.3
Sub3	1.287	-0.833	1.153	-0.778	-0.134	0.055	12.8	9.6
Sub4	1.244	-0.875	1.303	-0.922	0.059	-0.047	10.4	17.0
Sub5	1.004	-0.704	1.206	-0.890	0.202	-0.186	27.1	12.0
Sub6	1.046	-1.095	1.056	-1.077	0.010	0.018	4.8	10.1
Sub7	1.032	-1.173	1.144	-1.117	0.112	0.056	12.6	19.2
Sub8	1.038	-1.332	1.126	-1.397	0.088	-0.065	19.9	15.0
Sub9	1.534	-0.848	1.523	-0.806	-0.011	0.042	2.3	34.5
Sub10	1.073	-1.037	1.077	-1.040	0.004	-0.003	4.7	25.7
Sub11	1.295	-1.080	1.152	-0.997	-0.143	0.083	13.4	14.1
Sub12	1.064	-1.101	1.115	-1.127	0.051	-0.026	7.2	10.5
Sub13	1.048	-0.928	1.054	-1.008	0.006	-0.080	6.1	17.8
Sub14	1.35	-0.934	1.352	-0.935	0.002	-0.001	0.6	6.3
Sub15	1.201	-1.101	1.197	-1.078	-0.004	0.023	7.5	15.4
Sub16	1.235	-0.855	1.230	-0.908	-0.005	-0.053	1.4	17.5
Sub17	0.853	-1.135	1.096	-1.210	0.243	-0.075	20.7	18.3
Sub18	1.024	-1.251	1.042	-1.131	0.018	0.120	6.3	18.4
Sub19	1.196	-0.962	1.144	-0.930	-0.052	0.032	7.6	16.9
Sub20	1.171	-0.834	1.165	-0.899	-0.006	-0.065	2.1	20.0

Table 2: The bias and standard error (SE), defined as the square root of mean-squared error, for population (Pop) and individual dynamic parameters with 40 replications. The percentage of SE is based on the absolute value of the true parameter.

	$\log \phi$		$\log c$		$\log \delta$		$\log \lambda$		$\log \rho$		$\log N$		$\log k$	
	Bias	SE(%)	Bias	SE(%)	Bias	SE(%)	Bias	SE(%)	Bias	SE(%)	Bias	SE(%)	Bias	SE(%)
Pop	0.023	15.3	0.014	11.1	-0.008	13.1	-0.009	4.3	-0.010	9.1	-0.011	7.6	0.001	1.5
Sub1	0.284	21.1	-0.147	32.8	0.156	19.9	0.036	7.1	0.069	8.1	0.077	7.0	-0.242	4.1
Sub2	0.331	14.7	0.125	32.3	0.109	20.6	0.080	6.1	0.174	9.7	-0.112	5.8	0.089	1.6
Sub3	-0.121	11.0	-0.208	37.4	0.080	9.4	-0.073	5.5	0.052	5.2	-0.037	5.3	0.202	3.1
Sub4	-0.035	10.5	-0.206	31.1	-0.002	12.6	0.073	3.6	0.098	7.4	-0.144	5.2	-0.053	2.2
Sub5	0.125	24.3	-0.198	39.7	-0.048	23.3	0.010	6.1	-0.103	13.0	-0.116	9.1	-0.426	6.8
Sub6	-0.019	16.1	0.170	33.5	-0.061	16.9	0.053	4.0	0.236	10.6	0.154	5.5	0.085	3.2
Sub7	0.097	11.6	0.159	38.4	0.073	18.1	-0.158	5.4	-0.259	13.4	-0.171	3.9	0.243	3.4
Sub8	0.256	12.6	0.076	10.1	-0.128	19.3	-0.022	5.0	-0.234	13.8	-0.047	6.0	-0.141	2.4
Sub9	-0.108	9.0	0.027	31.9	0.063	22.8	-0.015	4.1	-0.244	12.8	0.021	4.2	0.001	1.3
Sub10	0.080	10.2	0.151	35.6	-0.108	33.5	0.118	4.2	0.185	8.4	0.079	3.3	0.041	2.1
Sub11	0.012	7.2	-0.075	22.9	0.010	16.5	0.075	6.8	0.194	9.3	-0.303	6.1	0.032	2.1
Sub12	0.269	31.1	-0.163	17.4	0.106	12.5	0.078	8.8	-0.117	15.2	-0.213	10.3	0.016	6.1
Sub13	0.153	12.6	0.093	17.2	-0.231	44.0	0.122	3.7	-0.120	7.4	-0.043	2.9	-0.013	2.2
Sub14	-0.053	9.8	0.063	27.1	0.049	16.3	0.103	4.8	0.159	8.7	0.085	4.8	0.028	2.2
Sub15	0.145	13.9	0.079	39.1	0.041	18.3	0.107	3.7	0.131	7.3	-0.008	4.9	-0.042	1.5
Sub16	0.044	5.3	-0.355	35.7	0.118	15.9	-0.103	4.3	0.117	7.0	-0.199	4.0	0.001	0.8
Sub17	0.295	20.7	-0.025	30.2	0.076	19.7	0.022	7.5	-0.120	10.1	-0.009	7.7	-0.185	4.0
Sub18	0.178	13.2	0.134	17.2	-0.128	32.5	0.152	5.4	0.130	6.5	0.183	4.7	-0.023	1.9
Sub19	0.047	9.6	-0.036	33.0	0.059	18.0	-0.065	4.9	0.218	10.1	-0.016	5.2	0.114	2.5
Sub20	0.177	11.8	-0.189	23.9	0.159	26.2	-0.061	5.3	0.091	5.8	-0.214	7.5	-0.017	2.0

Table 3: A summary of the estimated posterior means (PM) of population parameters and the corresponding 95% equal-tail credible intervals, where L_{CI} and R_{CI} denote the left and right credible limits of 95% credible intervals.

	ϕ	c	δ	λ	ρ	N	k
PM	24.9	3.06	0.37	98.1	0.081	975.6	0.000017
L_{CI}	10.6	2.79	0.33	89.1	0.073	886.3	0.000016
R_{CI}	57.7	3.37	0.41	107.9	0.089	1074.2	0.000018

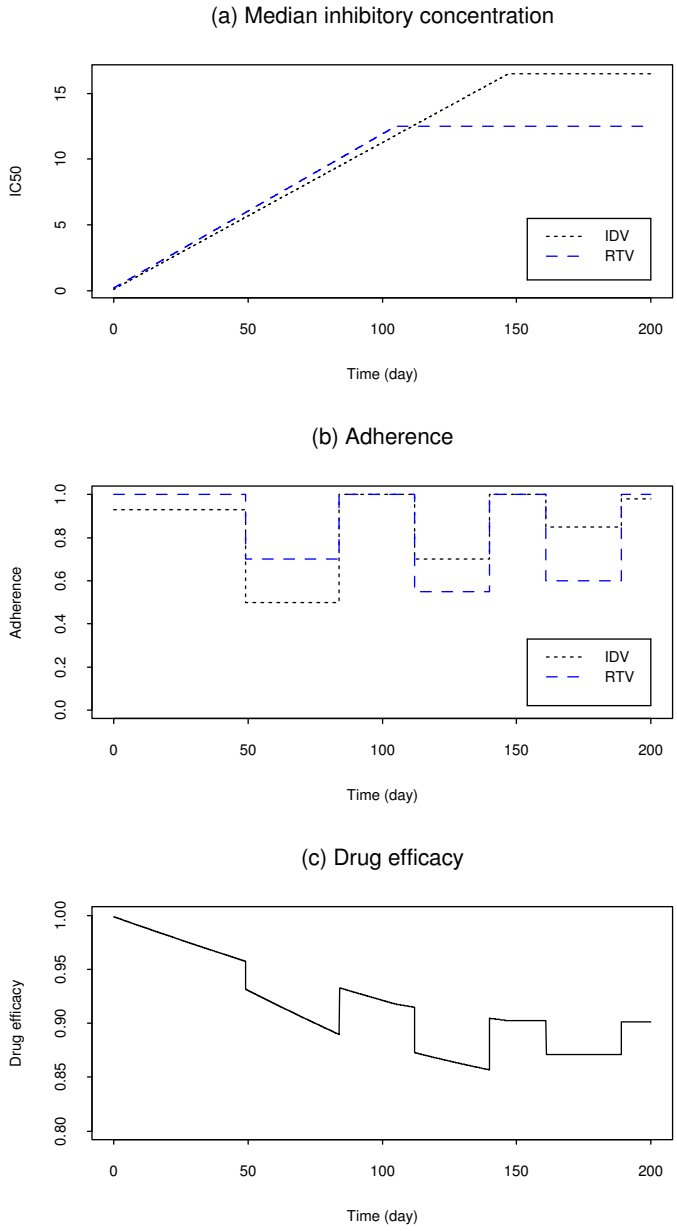

Figure 1: (a) The median inhibitory concentration curve $[IC_{50}(t)]$; (b) the time-course of adherence $[A(t)]$; (c) the time-course of drug efficacy $[\gamma(t)]$.

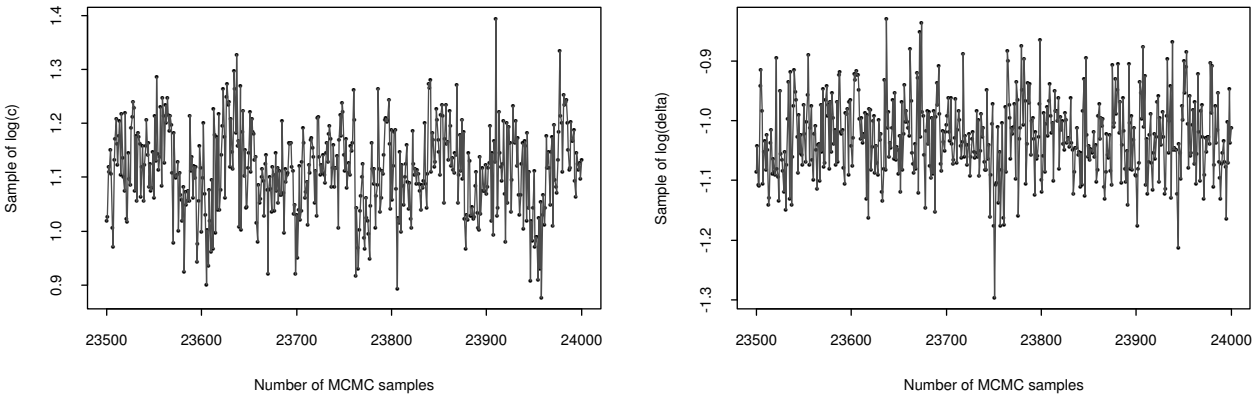

Figure 2: Trace plot to informally check parameter identifiability based on the last 500 samples drawn from the MCMC sampling scheme in Example 1.

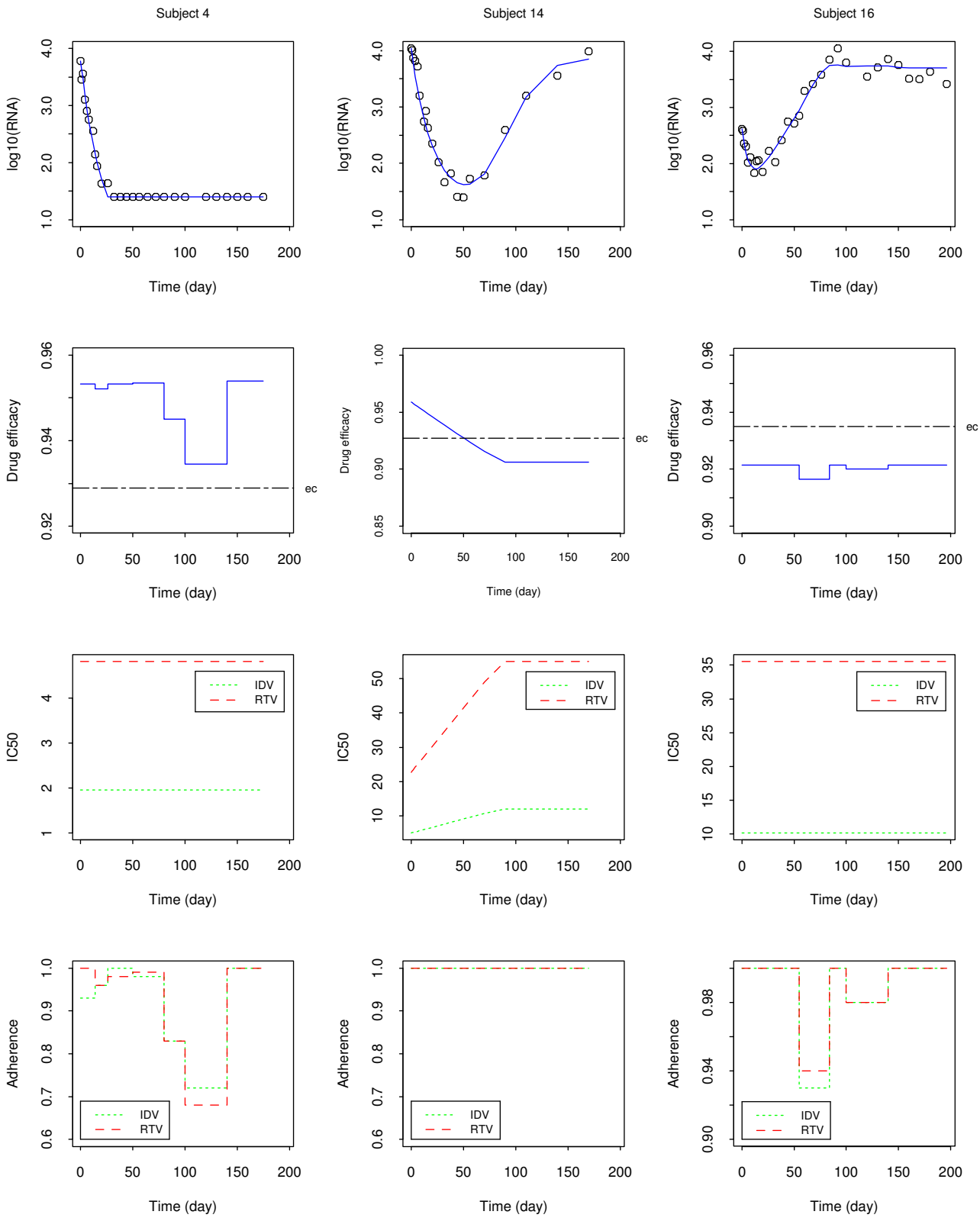

Figure 3: Individual fitted curves with generated viral load data in $\log_{10}$ scale, drug efficacy with threshold (e_c), as well as $IC_{50}(t)$ and adherence of the two PI drugs (IDV and RTV) for the three representative subjects from Example 1.

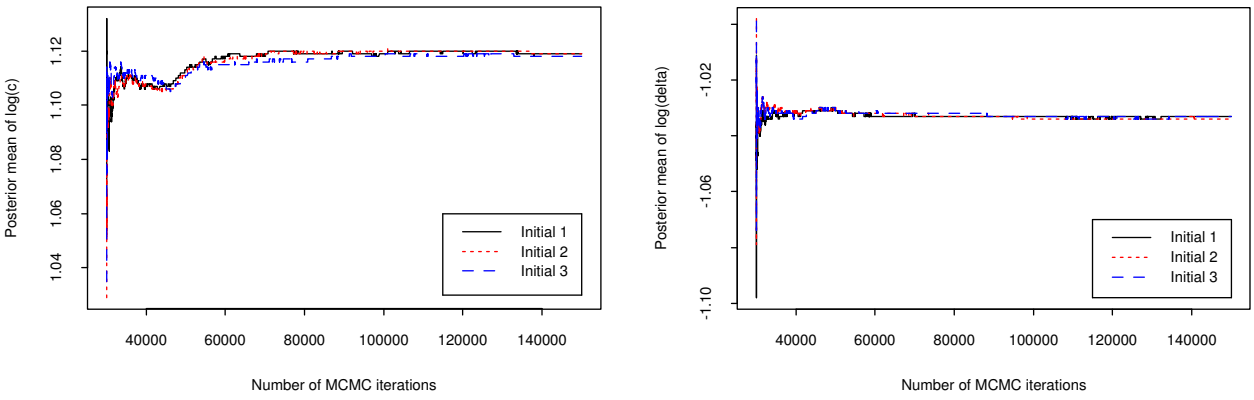

Figure 4: The number of MCMC iterations and convergence diagnostics with respect to three different initial values for Example 1.

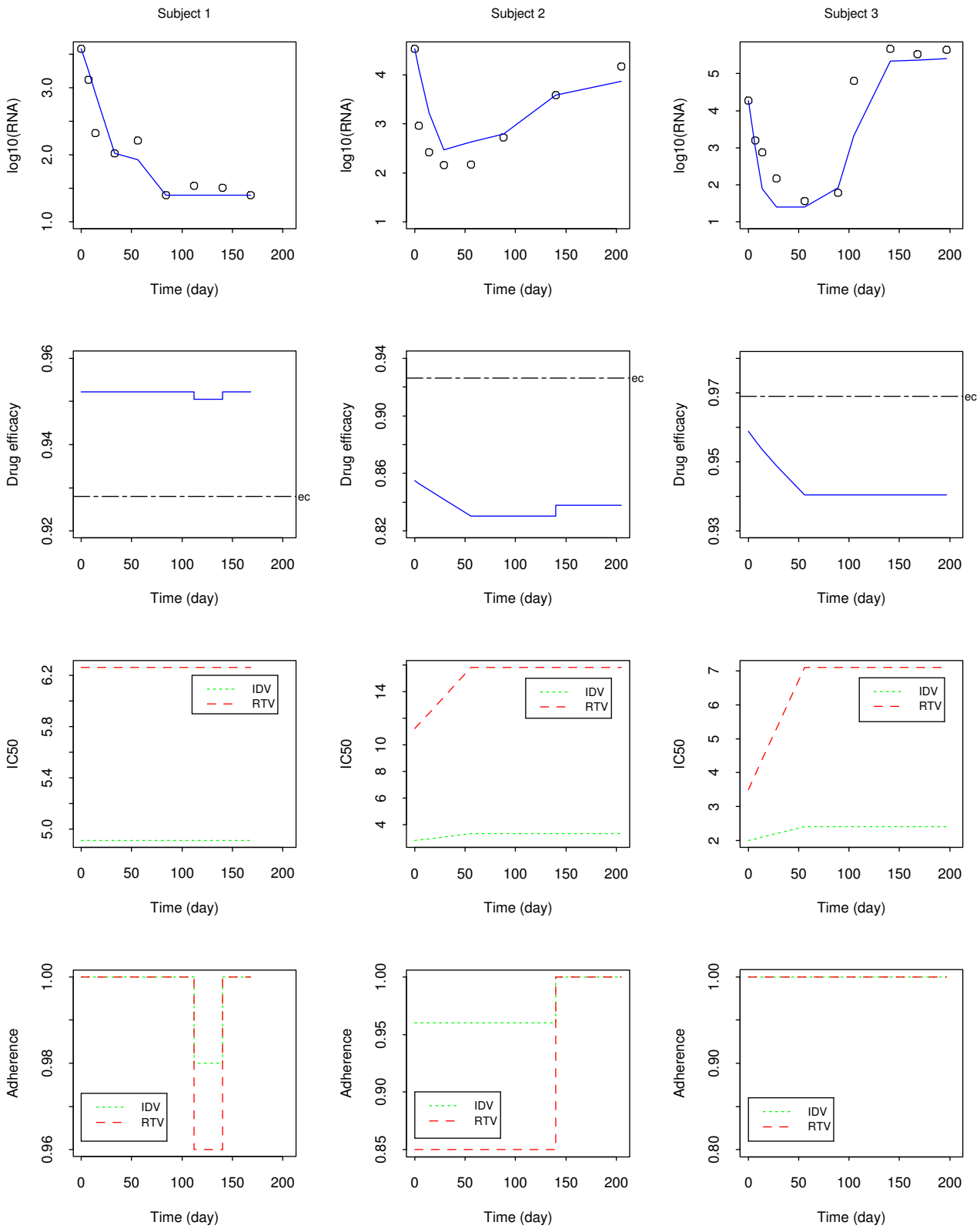

Figure 5: Individual fitted curves with observed viral load measurements in $\log_{10}$ scale, drug efficacy with threshold (e_c), as well as $IC_{50}(t)$ and adherence of the two PI drugs (IDV and RTV) for the three representative subjects.