# HIV MODEL PARAMETER ESTIMATES FROM INTERRUPTION TRIAL DATA

by

Rutao Luo

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Electrical & Computer Engineering

Fall 2012

© 2012 Rutao Luo
All Rights Reserved

**HIV MODEL PARAMETER ESTIMATES FROM INTERRUPTION TRIAL DATA**

by

Rutao Luo

Approved: _____

Ryan Zurakowski, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Kenneth E. Barner, Ph.D.
Chair of the Department of Electrical & Computer Engineering

Approved: _____

Babatunde Ogunnaike, Ph.D.
Interim Dean of the College of Engineering

Approved: _____

Charles G. Riordan, Ph.D.
Vice Provost for Graduate and Professional Education

## ACKNOWLEDGMENTS

I would like to acknowledge the advice and guidance of Dr. Ryan Zurakowski, my advisor and Dr. Michael Piovoso, our collaborator. Without their guidance, I would not have finished this degree. I also thank the members of Dr. Ryan Zurakowski's lab for their helpful discussion and suggestions.

I would like to thank my family members, especially my wife, Jing Wu, and my daughter Hannah Luo for supporting and encouraging me to pursue this degree.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## ABSTRACT

Mathematical models based on ordinary differential equations (ODE) have had significant impact on understanding HIV disease dynamics and optimizing patient treatment. A model that characterizes the essential disease dynamics can be used for prediction only if the model parameters are identifiable from clinical data. Most previous studies involved in parameter identification for HIV have used sparse data from the decay phase following the introduction of therapy. In this thesis, model parameters are identified from frequently sampled viral-load data taken from ten patients enrolled in the previously published AutoVac HAART interruption study, providing between 69 and 114 viral load measurements from 3-5 phases of viral decay and rebound for each patient. This dataset is considerablely larger than those used in previously published parameter estimation studies. Furthermore, the measurements come from two separate experimental conditions, which allows for the direct estimation of drug efficacy and reservoir contribution rates, two parameters that cannot be identified from decay-phase data alone. A Markov-Chain Monte-Carlo method is used to estimate the model parameter values, with initial estimates obtained using nonlinear least-squares methods. The posterior distributions of the parameter estimates are reported and compared for all patients.

# Chapter 1

# INTRODUCTION

## 1.1 HIV and Its Treatment

Human Immunodeficiency Virus (HIV) is a small, $\sim$ 10 kBp retrovirus that infects human CD4+ T cells and macrophages. Chronic, untreated infection by HIV results in a slow decline in functional CD4+ T cell counts. The depletion of these cells eventually leave them unable to support effective immune responses, leaving the host open to infection by any number of secondary infections. This state of severe induced immunodeficiency is termed Acquired Immunodeficiency Syndrome (AIDS).

By the most recent estimates of the Centers for Disease Control, there are approximately 1.1 million adults and adolescents in the United States infected with HIV, with approximately 56,000 new infections per year [10].

Modern HIV therapy uses a cocktail of drugs known as HAART (highly active antiretroviral therapy) to durably suppress viral replication. The components of HAART fall into five distinct classes of drugs, each with unique action and non-overlapping resistance mutation profiles.

The first class of drug is the Nucleoside/Nucleotide Analogue Reverse-Transcriptase Inhibitors (NRTIs). These drugs mimic amino acids or their precursors and are preferentially taken up by the viral enzyme reverse transcriptase, resulting in nonfunctional transcripts. There are seven drugs of this class currently approved for HIV treatment. Resistance mutations against this class tend to carry mixed levels of fitness cost, and there is moderate incidence of cross-resistance [53]. Multiple accumulated mutations are necessary to confer broad-class resistance against the NRTIs.

The second class of drugs are the non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs). These drugs bind to the reverse transcriptase (RT) enzyme in a

manner which allosterically deactivates the RT active site. There are four drugs in this class approved for HIV treatment. Resistance mutations against this class tend to carry very little fitness cost, and a single base-pair substitution mutation is capable of conferring clinically significant resistance against the whole NNRTI class [14].

The third class of drugs are the Protease Inhibitors (PIs). HIV viral proteins are unique in that they are first transcribed in a large, nonfunctional superprotein that is cleaved into functional units by the viral enzyme protease. Protease Inhibitors interfere with the action of this enzyme, resulting in the production of noninfectious virus particles. There are nine drugs in this class currently approved for HIV treatment. Resistance mutations against this class carry mixed levels of fitness cost [31], [30]; however, the drug resistance conferred against this class is rarely total [14]. Multiple-site mutations are necessary for broad-class resistance.

The fourth class of drugs are the newly introduced Integrase Inhibitors (IIs). These drugs inhibit the activity of integrase, an enzyme which facilitates the integration of HIV DNA into the host genome. Only one drug in this class has been approved for use in HIV therapy. This drug is only used as rescue therapy for patients who have developed significant resistance to all available drugs in the first three classes [14].

The fifth class of drugs are the Fusion Inhibitors (FIs). These drugs interfere with the fusion of the virus particle and the host cell by binding surface proteins on the host cell. Two drugs are available in this class. Both are expensive, and one is available through injection only. Both are used only as rescue therapy for patients with significant resistant to all available drugs in the first three classes [14].

The development of multi-drug regimens for HIV therapy has resulted in HIV infection becoming a chronic, manageable disease in first world countries [12] . The widespread use of three-drug regimens, usually consisting of two nucleoside/nucleotide analog reverse-transcriptase inhibitors (NRTIs) and either a non-nucleotide/nucleoside analog reverse-transcriptase inhibitor (NNRTI) or a boosted protease inhibitor (PI) usually provides adequate viral suppression and mutational barrier to maintain viral load at or below the measurement threshold indefinitely. The necessity of a three-drug

regimen, where each drug in the regimen targets separate viral epitopes, is due to the extremely high replication and mutation rates characteristic of HIV infection [16]. These make the evolution of viral strains resistant to a single drug inevitable. Three drugs, however, present a mutational barrier high enough to make such an evolutionary occurrence unlikely [41, 42]. While these three- drug regimens, known as highly active antiretroviral therapy, or HAART, are highly effective at suppressing the virus in the long term, some patients nevertheless experience viral load rebound, driven by the emergence of a viral mutant resistant to all three components of their HAART regimen.

## 1.2 Mathematical Model for HIV Dynamics

Beginning in the 1990's, researchers have analyzed the dynamics of human immunodeficiency virus (HIV) using nonlinear ordinary differential equation (ODE) models [2, 15, 33, 38, 54]. Using various mathematical models, they sought to simulate the dynamics of the virus or to help design a treatment. Many of these studies have attempted to identify model parameters from patient data, but the sparsity of measurements resulted in very large confidence intervals, especially in those cases in which only the viral load is measured. Furthermore, the experiment data that they used only included the period of viral decay following the introduction of therapy. Identification of the HIV model parameters under these conditions require an assumption that the drug efficacy is known, which in turn affect the estimates of the remaining parameters.

There are two problems for model identification without full-state information. The first is that the nonlinear ordinary differential equations used to model the dynamics of HIV virus have no closed-form solutions and the second is that the amount of data normally available is sparse. Typically, a patient during therapy routinely has his viral load tested every every 3 or 4 months, a rate too slow to accurately capture the dynamic characteristics.

In this thesis, the data used to identify model parameters are from the AutoVac study at IrsiCaixa HIV research foundation in Barcelona [46]. In the AutoVac study, 12 patients underwent a series of about 30-day treatment interruptions, followed by

resumption of suppressive therapy, and viral load measurements were taken at 3-day intervals during the interruption. This resulted in between 69 and 114 viral load measurements per patient, with between 38 and 77 data points per patient above the limit of detection. Therefore, data gathered from this particular experiment are sufficiently rich for model identification. Patients enrolled in the AutoVac study had all maintained undetectable viral loads on therapy for at least two years prior to the study, and all patients had baseline CD4+ T-Cell counts over 700. Patient nine and twelve both had interruptions without viral load rebound, and were excluded from this study due to insufficient usable data.

For HIV models, Stafford et al [52] have studied the identifiability of a 3 state model. Xia and Moog [58] analyzed the theoretical identifiability of a 4 state model and determined the minimal number of state measurements needed for estimating all model parameters. Frequently, only viral load data are available and in this case, not all parameters can be identified independently [52] [56]. Recently, Miao et al [32] also investigated some identifiability issues for viral dynamics.

For identifying the parameters of a viral dynamics model, two major methods are commonly used: nonlinear least squares [21] and Bayesian estimation [17, 18, 20, 39, 55]. In this thesis, we employ a Bayesian Markov-Chain Monte Carlo technique as in [39, 55], with nonlinear least-squares used to generate initial conditions for the MCMC technique. The primary difference between this work and previous works is the quality of the data used for estimation. The data used in [39] consists of 10 measurements from 12 patients taken at 5 time points. The data used in [55] consists of seven measurements from 42 patients taken at seven time points. The data used in [20] consists of 9 viral load measurements from 42 patients taken at nine time points, plus a single baseline measurement of phenotypic drug susceptibility and survey data on patient adherence. The data used in [28] used 18 measurements of viral load from 18 time points. All four of these studies only included data from a single virus decay phase following treatment initiation. As a result of the sparse data, these previous studies had to make a number of simplifying assumptions about parameter values in

order to preserve identifiability. By contrast, the AutoVac patient study provides us with between 69 and 114 viral load measurements from 10 patients from between 3 to 5 treatment interruptions cycles per patient. The high quality of the data used in this study allows reliable estimation of parameter values without resorting to the simplifying assumptions used in previous studies.

We consider the following mathematical model characterizing the viral dynamics for a patient:

$$\begin{aligned} \dot{x}(t) &= \lambda - dx(t) - \beta(1-\eta u)x(t)v(t) \\ \dot{y}(t) &= \beta(1-\eta u)x(t)v(t) - ay(t) + \lambda_y(t) \\ \dot{v}(t) &= \gamma y(t) - \omega v(t) \end{aligned} \tag{1.1}$$

There are three states: $x$, the concentration of target CD4+ T cells; $y$, the concentration of actively infected CD4+ T cells; $v$, the viral load. $\lambda$ is the proliferation rate and $d$ is the death rate of target CD4+ T cells; $\beta$ is the infection rate; $\eta$ is the drug efficacy; $a$ is the death rate of actively infected cells; $\lambda_y(t)$ is the contribution of the reservoir to actively infected CD4+ T cells; $\gamma$ is the rate of free virus production by infected cells; $\omega$ is the clearance rate for the free virus. The drug application $u$ is 0 during interruptions and 1 during treatment. This is a variation of a model first proposed in [36], with the addition of the $\lambda_y(t)$ term describing the additional contribution of infected cells from all viral reservoir processes. This model is essentially the same as the model identified against patient data in the previous studies [17, 18, 20, 21, 55].

Highly active antiretroviral therapy (HAART) has proven effective to reduce the active viral load [13, 37] and is standard care for HIV patients. However, it cannot eradicate the virus completely [6, 35]. Although scientists suspect that the existence of long-term latent reservoirs in patients is the main reason for viral persistence [5, 9], there is little quantitative understanding of their contribution, mainly because of the difficulty measuring the virus reservoir directly. Some HIV investigators have proposed mathematical models to describe the dynamics of long-term latent reservoirs [24, 44, 45]. Little research has been done to estimate the parameters of these models quantitatively based on clinical data. In this study, the total contribution of the reservoir processes

to the actively infected CD4+T cells is estimated from standard viral-load time-series data. We analyze the identifiability for each parameter in Model 1.1 using differential algebra tools. The implementation of Bayesian estimation method is presented, and the results of the Bayesian estimation method are reported. This work has been published in PloS One and a shorter version of this work can be found in [26].

# Chapter 2

# METHODS

## 2.1 Experimental Methods

The previously published clinical study [46] was carried out in accordance with a human subjects protocol approved by the institutional ethics review committee at the University Hospital Germans Trias i Pujol in Barcelona, Spain. Written informed consent was obtained from all study participants. De-identified patient data was shared in accordance with a protocol approved by the University of Delaware Institutional Review Board.

This research described here uses data from a previously published study. The measurements which are the focus of this work have been previously described in [46]. Briefly, a randomized prospective Structured Treatment Interruption study enrolled 26 HIV-1 positive asymptomatic adults with no detectable virus for at least two years prior to entering the study (limit of detection 50 virions per ml). Fourteen were randomized to a control group, continuing their previous HAART regimens. Twelve were randomized to the experimental group, and underwent between three and five cycles of interrupted antiviral therapy, remaining off therapy until two consecutive viral load measurements above 3000 virions/ mL were reached, or for a maximum of 30 days, then re-initiating the original HAART regimen for 90 days before the next interruption cycle began. HIV-1 RNA PCR quantitative analysis was performed on samples collected three times weekly following treatment interruption, and then weekly for the two months following re-initiation of treatment.

## 2.2 Modeling Persistence of Latent Reservoirs

Although highly active antiretroviral therapy (HAART) efficiently suppress viral load under undetectable level, current regimens can not eradicate the virus completely [7, 11, 34, 45, 51]. One possible reason is the persistent replication of HIV at a very low level, even under HAART conditions [1, 45, 49]. Another possible reason is the existence of stable reservoirs of latently infected cells [3, 4, 45]. These two possibilities are not mutually exclusive, and it is likely that a combination of persistent viremia and reservoir activation combine to maintain the reservoirs [45].

Rong and Perelson proposed models to describe the dynamics of the latent reservoir [45]. However, in order to make the model identifiable from viral-load data, the model must be simplified. Siliciano et al. [50] found that the average half-life of the latent reservoir in resting $CD4^+T$ cells is 44 months, which means it is extremely stable. There is no strong evidence that the activation rate of the reservoir is constant, however, and a combination of various factors may contribute to the maintenance of a residual virus load during effective antiviral suppression. Therefore, in Equation 1.1 $\lambda_y(t)$ represents the total average contribution of reservoir dynamics to the active infected cell compartment $y$ during the treatment period. The time between ceasing antiviral therapy and the viral load reaching measurable levels (the rebound time) is very sensitive to the value of $\lambda_y(t)$, and consequently the goodness of fit for the entire model is also very sensitive to $\lambda_y$. Surprisingly, a single constant value for $\lambda_y$ provided an excellent fit for all ten patients for all interruption cycles, implying that the average contribution of reservoir dynamics to the active infected cell compartment is relatively constant between several interruption cycles.

## 2.3 Identifiability Analysis

Equation 1.1 is a special case of the following general nonlinear model:

$$\dot{X}(P,t) \;=\; F(X(P,t),P) \tag{2.1}$$

where, in our case, $X(P,t)$ is the state vector $[x,y,v]'$ in our study, $P$ is the parameter vector

$[\lambda, d, \beta, u, a, \lambda_y, \gamma, \omega]'$, $F$ is the function which describes system dynamics.

We adapt the concepts of identifiability from [47] [22].

**Definition 1.** Equation 2.1 is said to be globally identifiable from the given states if the equation $F(X(P,t),P) = F(X(P^*,t),P^*)$ has only one solution $P = P^*$.

**Definition 2.** Equation 2.1 is said to be locally identifiable from the given states if in some open neighborhood, $U_{p^*}$, around the true parameter vector, the equation $F(X(P,t),P) = F(X(P^*,t),P^*)$ has only one solution $P = P^*$ and $P^* \in U_{p^*}$.

The identifiability of HIV dynamic models has been analyzed previously [32] [57]; however, these previous works assumed more than one state could be measured. Here, only the viral load data are assumed to be available. Differential algebra is used to analyze the identifiability issue of Model 1.1. Details of differential algebra are found in [43] [48] [47]. The steps in this analysis are as follows:

i) For Equation 1.1, in order to generate the differential polynomial, we choose the following order relation: $v < \dot{v} < \ddot{v} < \dddot{v}$.

ii) Based on the above order, the normalized characteristic polynomial of $v$ is calculated:

$$\begin{aligned} A &= v\dddot{v} - \dot{v}\ddot{v} + \beta(1-\eta u)(v)^2\ddot{v} + (a+d+\omega)v\ddot{v} - (a+\omega)(\dot{v})^2 + \\ &\quad (a\beta(1-\eta u) + \beta\omega(1-\eta u))(v)^2\dot{v} + (ad+d\omega)v\dot{v} + \lambda_y\dot{v} + \\ &\quad a\beta\omega(1-\eta u)(v)^3 + (ad\omega - \beta\lambda_y(1-\eta u) - \beta\gamma\lambda(1-\eta u))(v)^2 - d\lambda_y v \\ &= 0 \end{aligned} \tag{2.2}$$

Equation 2.2 is generated by a) solving for $y$ from the last equation of Equation 1.1, b) substitute for $y$ in the middle equation of Equation 1.1, c) solve for $x$ from that equation and d) substitute for $x$ in the first equation of Equation 1.1. Knowledge of $v$ allows one to estimate the coefficients in Equation 2.2. This characteristic polynomial does not contain the states $x$, $y$ and their derivatives, but still describe the viral dynamics of Equation 1.1.

iii) By extracting the coefficients in the above polynomial and setting them as the estimated values, we obtain:

$$\begin{aligned}
\beta(1-\eta u) &= a_1 \\
a+d+\omega &= a_2 \\
a+\omega &= a_3 \\
a\beta(1-\eta u)+\beta\omega(1-\eta u) &= a_4 \\
ad+d\omega &= a_5 \\
\lambda_{\mathrm{y}} &= a_6 \\
a\beta\omega(1-\eta u) &= a_7 \\
ad\omega-\beta\lambda_{\mathrm{y}}(1-\eta u)-\beta\gamma\lambda(1-\eta u) &= a_8 \\
d\lambda_{\mathrm{y}} &= a_9
\end{aligned} \tag{2.3}$$

iv) The identifiability of each parameter can be checked by checking the injectivity of the the map defined in Equation 2.3. All parameters except $\gamma$ and $\lambda$ are uniquely identifiable; the product $\gamma\lambda$ is uniquely identifiable. Since the patients enrolled in the AutoVac study had all maintained undetectable viral loads on therapy for at least two years prior to the study, it is reasonable to assume that the initial conditions of this study are the steady states in Equation 1.1. Therefore, the steady state of $x$ in Equation 1.1, $\frac{\lambda}{d}$, is set as the number of target cells at the time of beginning this study. Under this assumption, the value of $\lambda$ is able to be determined as the product of the initial number of target cells $x(0)$ and the estimated decay rate of target cell $d$.

Although theoretically $\omega$ is uniquely identifiable, the current best estimate of $\omega$ is between 9 and 36 $\frac{1}{day}$ [40]. Estimation of $\omega$ would require very high frequency measurements (several measurements per hour, much faster than our data). Therefore the value of $\omega$ is set as 18.8 $day^{-1}$, the mean of 1/2 life of virus estimated from [40], and the virus dynamics are treated as a singular perturbation to the system. The data from AutoVac study are gathered in two phase: treatment interruption and therapy resumption. During the treatment interruptions, the drug application $u$ is equal to 0. In the nonliear least-square method, the data gathered during treatment interruptions is used for estimating the globally identifiable parameters, $\lambda, d, \beta, a, \lambda_y, \gamma$. Similarly,

drug efficacy, $\eta$, is also globally identifiable and can be estimated from the data gathered during treatment resumption. During resumption of treatment, $\lambda, d, \beta, a, \lambda_y, \gamma, \omega$ are fixed as the values that were estimated during interruption and only the drug efficacy, $\eta$, is identified.

### 2.4 Bayesian Estimation

During the AutoVac study treatment for each patient was interrupted and after a period of time, restarted. This cycle of interruption and reinstating the treatment is repeated 3 to 5 times. Shown in Fig.2.1 is a plot of the logarithm of the viral load for a particular patient versus time in days. In order to generate an initial value for the MCMC method, a two-step least-squares method was used, using data from the first three interruption cycles. The data from the period in which the treatment is interrupted, region 1, is used to estimate the 5 parameters, $\lambda, d, \beta, a$ and $\gamma$ using a constrained nonlinear least squares method. The data from region 2, where treatment is reinstated, contains information about the drug efficacy. With the value of the six parameters fixed to those values found by least squares in region 1, the data of region 2 is used to estimate the drug efficacy.

From the steady-state values of Equation 1.1, the relationship between $d$ and $\lambda$ can be written as:

$$d = \frac{\mathbf{x}(0)}{\lambda} \tag{2.4}$$

.

Where $\mathbf{x}(0)$ is the initial measurement of CD4+ T cells taken for each patient at the beginning of the study (as in [39]). Therefore, in this method, the parameter set $[\lambda, \beta, a, \gamma, \lambda_y, \eta]^T$ are estimated. $\eta$ is calculated for each iteration by least-squares subject to the values of the other five parameters. The MCMC approach taken here is based on the Metropolis-Hasting algorithm [18]. Assume that the $i^{th}$ subject, we have

$m_i$ measurements of viral load. We denote the parameters as:

$$\mu = [\log(\lambda), \log(\beta), \log(a), \log(\gamma), \log(\lambda_y)]^T$$
$$\theta_i = [\log(\lambda_i), \log(\beta_i), \log(a_i), \log(\gamma_i), \log(\lambda_{yi})]^T$$
$$V = \{V_{ij}(\theta_i, t_j), i = 1, \cdots, n; j = 1, \cdots, m_i\}$$
$$Y = \{y_{ij}, i = 1, \cdots, n; j = 1, \cdots, m_i\}$$

The logarithm is used to ensure that all estimates of the parameters are positive. The vectors, $\mu$ and $\theta_i$ are the logarithm of the parameters for the population level and the logarithm of the parameters for the $i^{th}$ individual respectively. The initial values of $\mu$ and $\theta_i$ are set as the results of Patient 2 from the nonlinear least-square method, which are $[3.5877, -12.6082, -1.7614, 8.3551, -10.1926]^T$. The matrix $Y$ is the matrix of the logarithm of available measurements to base 10 for all the patients. Let $V(t)$ denote the solution of the differential equation and $V_{ij}(\theta_i, t_j)$ is the value of $V(t)$ for the $i^{th}$ patient using parameters $\theta_i$ at time $t_j$.

Following the iterative MCMC algorithm of [18] [19], the implementation steps can be written as:

1.Initialize the chain with initial values $(\sigma^{-2(0)}, \mu^{(0)}, \Xi^{-1(0)}, \theta_i^{(0)})$.

2.Use Gibbs sampling steps to update $\sigma^{-2}$,$\mu$, and $\Xi^{-1}$ and use Metropolis-Hastings algorithm to update $\theta_i$:

(a) $\sigma^{-2(k)} \sim Ga\left(a + \frac{m_i}{2}, A^{-1(k)} | \{\theta_i^{(k-1)}, \Xi^{-1(k-1)}, Y\}\right)$;

$\mu^k \sim N\left(B^{-1(k)} C^k, B^{-1(k)} | \{\Xi^{-1(k-1)}, \sigma^{-2(k)}, Y\}\right)$

$\Xi^{-1(k)} \sim Wi\left(D^k - 1, 1 + \nu | \{\sigma^{-2(k)}, \mu^k, Y\}\right)$

where $A^k = b^{-1} + \frac{1}{2} \sum_{j=1}^{m_i} \left| y_{ij} - log10\left(V_{ij}\left(\theta_i, t_j\right)\right)\right|^2$, $B^k = \Xi^{-1(k)} + \Lambda^{-1}$, $C^k = \Xi^{-1(k-1)} \theta_i^{(k-1)} + \Lambda^{-1} \eta$ and $\mathrm{D} = \Omega^{-1} + \left(\theta_i^{(k-1)} - \mu^{k-1}\right)\left(\theta_i^{(k-1)} - \mu^{k-1}\right)^T$.

$Ga$ is the gamma distribution and $Wi$ is the Wishart prior distribution.

All the hyper-parameters, $a, b, \eta, \Lambda, \Omega$ and $\nu$ are known as follows:

$$
\begin{aligned}
&a = 4.5;\\
&b = 9;\\
&\eta = [3.5877, -12.6082, -1.7614, 8.3551, -10.1926]^{\mathrm{T}};\\
&\Lambda = diag([100, 0.2, 0.4, 1000, 0.3]);\\
&\Omega = diag([2.5, 2, 2, 2.5, 2]);\\
&\upsilon = 8
\end{aligned}
$$

the hyper-prior values for the initial variance of $\theta_i$, $\Lambda$, are sufficiently large that this may be considered a non-informative prior distribution. For comparison, the analysis in [39] had initial variance for $\lambda$ and $\gamma$ of 0.12 and 0.0594 respectively, heavily biasing the posteriors to the priors, and the analysis in [20] had initial variance for the parameters $\lambda, d, \beta, a, \gamma, \omega$ of 0.005, also significantly biasing the posterior distributions of these parameter estimates to the prior distributions.

(b) Generate a new value $\phi$ for $\theta_i^{(j)}$ from the proposal prior distribution from $\theta_i^{(j-1)}$ $\left(\theta_i^{(j)} \sim N\left(\theta_i^{(j-1)}, \Lambda\right)\right)$. Evaluate the acceptance probability of this move by applying Metropolis-Hastings algorithm. If this move is accepted, $\theta_i^{(j)} = \phi$. If not, $\theta_i^{(j)} = \theta_i^{(j-1)}$.

3. Repeat step 2 until the chain converges.

To obtain reasonable results from the MCMC method, good initial estimates of $\theta_i$ are needed. The constrained least squares approach described previously is used to get an initial estimate. The parameters are constrained so that basic reproduction ratio during treatment interruptions, $R_0 = \frac{\lambda\beta\gamma}{da\omega}$ is greater than 1. If it were less than 1, the virus would eventually be eliminated.

The above procedure was applied to the data for 10 patients with sufficient data. The MCMC procedure produced 200,000 possible sets of parameters for each patient that are consistent with the patients' data. For the purposes of analysis, the first 50,000 iterations were discarded to allow the chain to converge, leaving 150,000 parameter sets

per patient for the final analysis. From this result, the marginal probability densities for of the six parameters can be established.

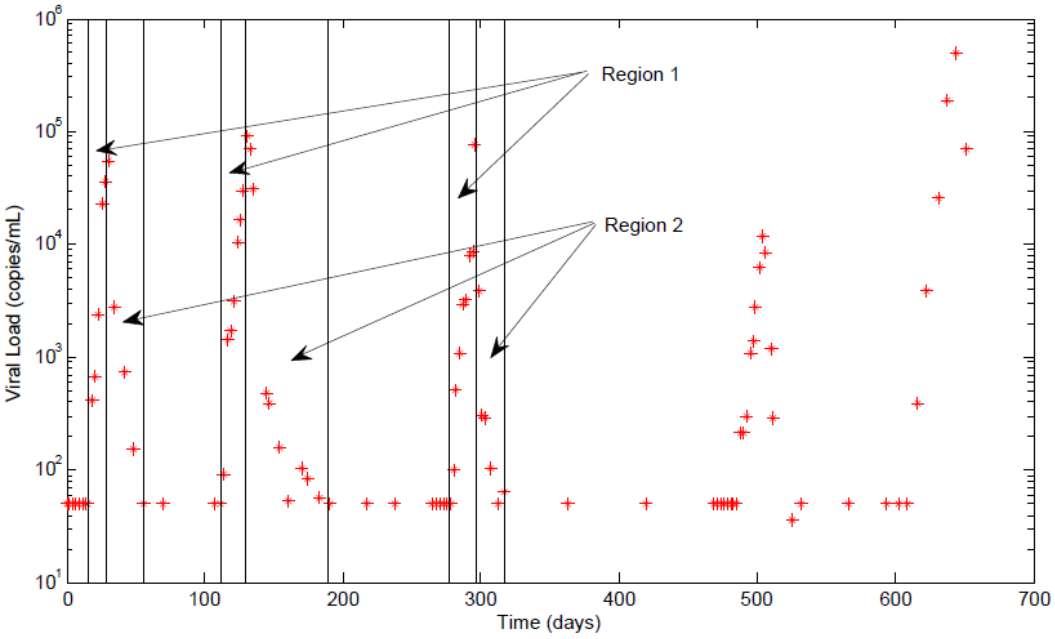


**Figure 2.1: The viral load data for Patient 1**

# Chapter 3

# RESULTS

## 3.1 Nonlinear Least Squares Estimation

Parameter estimates were generated for each of 10 patients using the nonlinear least squares method. Of the 12 patients in the study, 2 had no detectable virus after an interruption, leaving insufficient data above the measurement threshold to identify model parameters. Although the nonlinear least-square method can not guarantee to give globally optimal results, it can provide us good estimates for the prior distribution of the MCMC method. The results of Patient 2 by using this identification method are shown in Fig. 3.1.

## 3.2 Bayesian Estimation

The MCMC model fitting procedure was run for each of 10 patients with sufficient data. Histograms of the marginal posterior distributions for the six parameters for each of the 10 patients are shown in Fig.3.2.

Note that the parameters are not independent, and the parameter vectors should be considered as complete sets. Table.3.1 shows the average correlation coefficient among each different pair of parameters; it is clear that most pairs of parameters would be considered highly correlated. In addition to this first-order correlation between parameters, there are also higher order nonlinear correlations. Fig. 3.3 shows the high level of correlation between the product of the parameters which form the numerator and denominator of $R_0 = \frac{\lambda\beta\gamma}{da\omega}$, further emphasizing the need to consider parameter vectors rather than individual parameters. Values chosen independently from each parameter's distribution can easily generate a parameter set which is a particularly poor representation of the data. Consequently, we also report as tables in the supplementary

material the entire posterior distribution for each of the 10 patients (shown in supplementary tables S1-S10). The posterior distribution generated by this MCMC method provides a database for testing the robustness of treatment optimization strategies, such as those described in [23–25, 27].

**Table 3.1:** The average correlation coefficients between parameters for all ten patients.

| | $\log_{10}(\lambda)$ | $\log_{10}(\beta)$ | $\log_{10}(a)$ | $\log_{10}(\gamma)$ |
|---|---|---|---|---|
| $\log_{10}(\lambda)$ | 1 | 0.0953 | 0.6632 | 0.4278 |
| $\log_{10}(\beta)$ | 0.0953 | 1 | 0.1228 | -0.6421 |
| $\log_{10}(a)$ | 0.6632 | 0.1228 | 1 | 0.5817 |
| $\log_{10}(\gamma)$ | -0.6421 | 0.5817 | 0.6632 | 1 |

The histograms shown in Fig.3.2 demonstrate the range of values of each parameter, and Table 3.2 gives the maximum likelihood estimate for each parameter. This can be subtly misleading, because the parameters are correlated in some relationship, however, and the joint distributions of the parameters are represented by the data contained in the supplementary material. From the histograms it is clear that the distributions for the parameters $\beta$,$a$,and $\gamma$ vary little between patients, indicating that the infection rate and burst size of the virus and the death rate of infected cells may not vary much by patient. By contrast, the parameters $\lambda$, $\lambda_y$, and $\eta$ vary significantly between patients, indicating that the regeneration rate of CD4+ T cells, the reservoir contribution rate, and the drug efficacy may vary significantly by patient.

Table.3.3 gives a summary of the estimated population parameters (the average value of the 10 identified patients), compared with those from previously published papers. $R_{0pre}$ is the $R_0$ without treatment. $R_{0post}$ is the $R_0$ with treatment. The values for parameters $\lambda$,$d$,$\beta$,$a$, and $\gamma$ are consistent with the previously published best estimates for these parameters. In particular, the maximum likelihood estimates for the death rate of target cells $d$ ranged from $0.045-0.45\frac{1}{day}$, slightly faster than the $0.01\frac{1}{day}$ rate used in [45] but in perfect agreement with the estimates obtained from patient data in [18,20] $0.09-0.41\frac{1}{day}$ and $0.07-0.09\frac{1}{day}$ respectively). Our maximum likelihood

estimates for the death rate of infected cells $a$ ranged from $0.18 - 2.3\frac{1}{day}$, in agreement with the current best estimate of $\frac{0.7-1.3}{day}$ [28, 45]. Our maximum likelihood estimate of the density-dependent infection rate $\beta$ ranged from $2 \times 10^{-6} - 6 \times 10^{-6}\frac{mL}{virion \times day}$, compared to $1.6 \times 10^{-5} - 1.8 \times 10^{-5}\frac{mL}{virion \times day}$ [18] and $(1 \times 10^{-7} - 3 \times 10^{-7}\frac{mL}{virion \times day}$ [39]. Our maximum likelihood estimates of the target cell recruitment rate $\lambda$ ranged from $35 - 760\frac{cells}{\mu L \times day}$, higher than the comparable range of $86 - 111\frac{cells}{\mu L \times day}$ reported in [20]; however, this is expected, as the inclusion criteria for our experiment resulted in patients with much healthier immune systems overall compared to the patients in [20]. Our estimates of the virus production rate $\gamma$ range from $2.4 \times 10^3 - 9.8 \times 10^3 \frac{virions \times \mu L}{cells \times mL}$, higher than the $1 \times 10^2 - 1 \times 10^3 \frac{virions \times \mu L}{cells \times mL}$ range reported in [20]; however, this is due to the estimate in [20] of $\omega \approx 3\frac{1}{day}$, as opposed to our fixed estimate of $\omega = 18.8\frac{1}{day}$.

The maximum likelihood values of the reservoir contribution rate ranged over three orders of magnitude, from $2 \times 10^{-6}\frac{cells}{\mu L \times day}$ for Patient 1 to $1 \times 10^{-3}\frac{cells}{\mu L \times day}$ for Patients 3 and 10. This suggests that the replenishment of the active compartment by the viral reservoirs is very heterogenous between patients. However, it is noteworthy that a single constant value of $\lambda_y$ was able to accurately predict rebound time across multiple interruption cycles for the same patient despite variation in interruption length, indicating that the replenishment rate for a given patient is relatively constant over the course of the experiment.

The efficacy of the antiviral drugs is estimated directly from the viral load data. To our knowledge, this is the first time this has been done. Previous estimates of model parameters have typically inferred drug efficacy from PK/PD (Pharmacokinetic/Pharmacodynamic) data in the plasma, resulting in estimates of $\eta \approx 0.95$ [20], or relied heavily on the assumption that the first measurement was at steady-state [39], resulting in estimates between $0.64 \leq \eta \leq 0.84$. By contrast, our direct estimate of $\eta$ from the viral load data give us maximum likelihood estimates ranging from $\eta = 0.67$ to $\eta = 0.88$. This indicates that the estimates based on plasma pharmacokinetic data may overestimate the true drug efficacy, though the estimates which relied exclusively on the first measurement were consistent with our results.

Fig.3.4 shows the histograms of the coefficient of determination $R^2$ for each of the 10 patients. The statistical significance threshold for the fit relative to an average measurement model was calculated using the F-test ($P < 0.05$ is considered statistically significant); this shows that, for all patients except Patient 6, all 150,000 parameter sets in the posterior distribution would be considered a statistically significant fit to the data if considered in isolation.

The viral load fitted by the model using the maximum likelihood estimates of the parameters and simulation results for Patient 1 are shown in Fig.3.5.

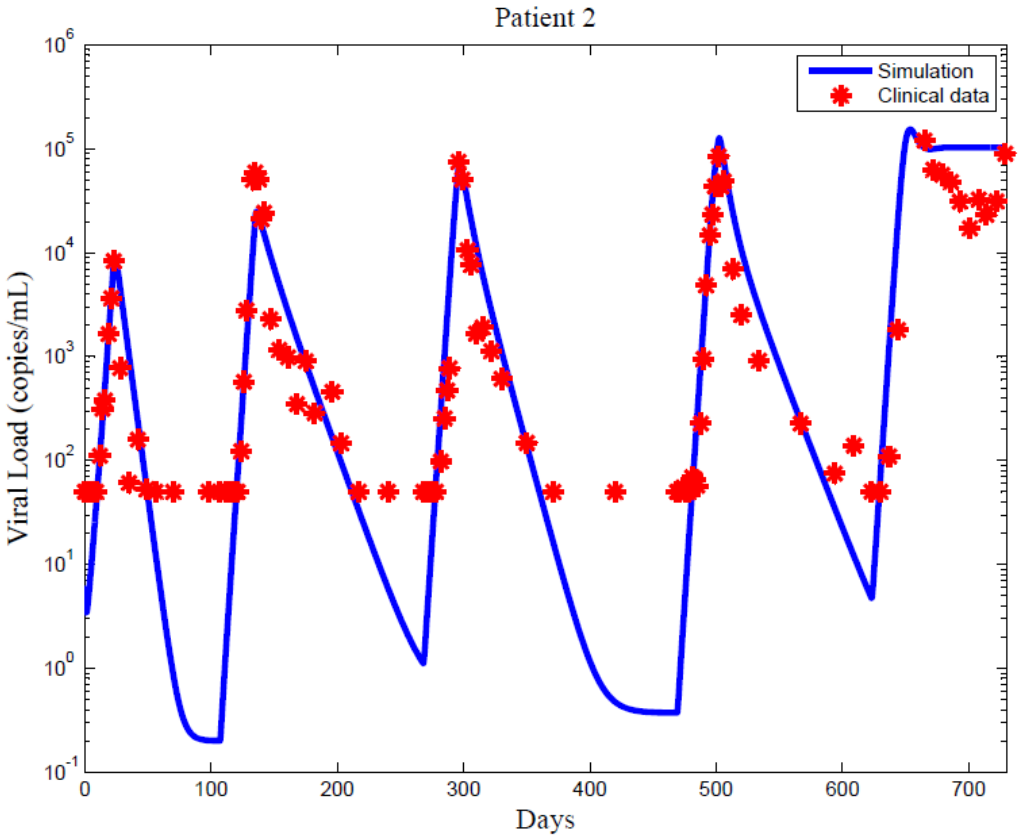


**Figure 3.1: Model fitting for identified patients.** Red star: experimental data (detection limit: 50 copies/mL); solid line: estimate.

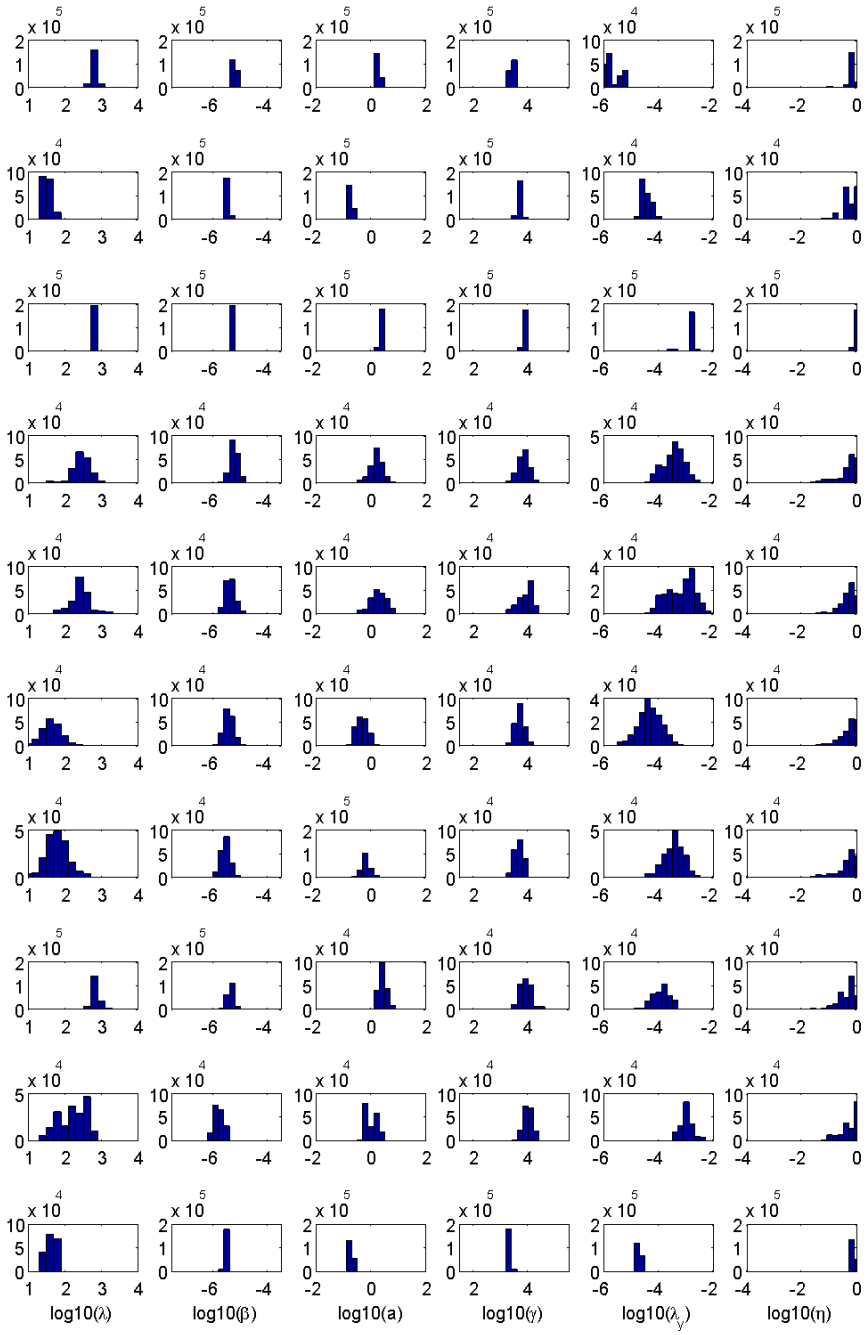


**Figure 3.2: The marginal posterior distribution of each parameter for ten different patients**

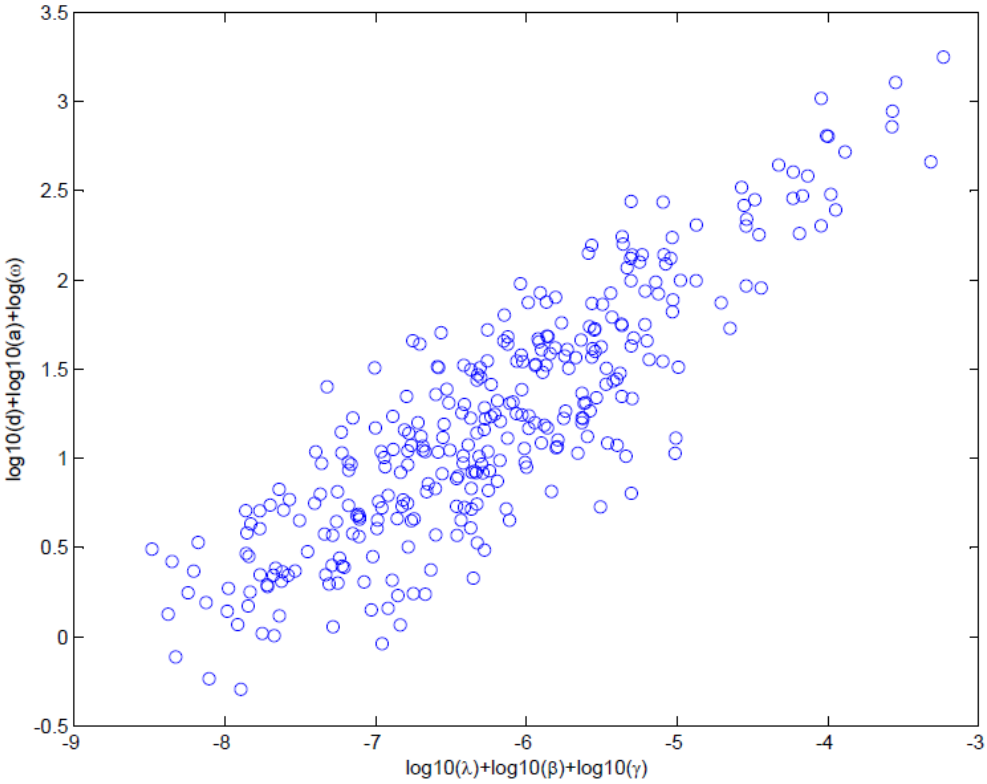


**Figure 3.3: Typical scatter plot of $R_0$ correlation** The heavy correlation between the elements in the numerator and denominator of $R_0$ demonstrates the strong higher-order correlation between parameters.

**Table 3.2:** Parameter identification results for each patient, reported as mean(standard deviation).

| Parameter | Unit | Patient 1 | Patient 2 |
|---|---|---|---|
| $\log_{10}(\lambda)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | 2.77(0.07) | 1.54(0.12) |
| $\log_{10}(d)$ | $\log_{10}(\frac{1}{day})$ | $-0.68(0.07)$ | -1.33(0.12) |
| $\log_{10}(\beta)$ | $\log_{10}(\frac{mL}{copies \times day})$ | -5.26(0.09) | -5.48(0.06) |
| $\log_{10}(a)$ | $\log_{10}(\frac{1}{day})$ | 0.23(0.08) | -0.76(0.07) |
| $\log_{10}(\gamma)$ | $\log_{10}(\frac{copies \times \mu L}{cells \times mL \times day})$ | 3.45(0.12) | 3.69(0.05) |
| $\log_{10}(\eta)$ | $--$ | -0.14(0.18) | -0.10(0.28) |
| $\log_{10}(\lambda_y)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | -5.67(0.30) | -4.43(0.17) |
| Parameter | Unit | Patient 3 | Patient 4 |
| $\log_{10}(\lambda)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | 2.88(0.04) | 2.45(0.27) |
| $\log_{10}(d)$ | $\log_{10}(\frac{1}{day})$ | -0.34(0.04) | -0.61(0.27) |
| $\log_{10}(\beta)$ | $\log_{10}(\frac{mL}{copies \times day})$ | -5.35(0.01) | -5.23(0.15) |
| $\log_{10}(a)$ | $\log_{10}(\frac{1}{day})$ | 0.37(0.04) | 0.19(0.24) |
| $\log_{10}(\gamma)$ | $\log_{10}(\frac{copies \times \mu L}{cells \times mL \times day})$ | 3.84(0.03) | 3.82(0.22) |
| $\log_{10}(\eta)$ | $--$ | -0.05(0.03) | -0.17(0.43) |
| $\log_{10}(\lambda_y)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | -2.94(0.20) | -3.45(0.38) |
| Parameter | Unit | Patient 5 | Patient 6 |
| $\log_{10}(\lambda)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | 2.43(0.27) | 1.64(0.29) |
| $\log_{10}(d)$ | $\log_{10}(\frac{1}{day})$ | -0.64(0.27) | -1.28(0.29) |
| $\log_{10}(\beta)$ | $\log_{10}(\frac{mL}{copies \times day})$ | -5.36(0.16) | -5.42(0.18) |
| $\log_{10}(a)$ | $\log_{10}(\frac{1}{day})$ | 0.26(0.28) | -0.32(0.21) |
| $\log_{10}(\gamma)$ | $\log_{10}(\frac{copies \times \mu L}{cells \times mL \times day})$ | 3.90(0.25) | 3.70(0.17) |
| $\log_{10}(\eta)$ | $--$ | -0.17(0.37) | -0.15(0.38) |
| $\log_{10}(\lambda_y)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | -3.18(0.51) | -4.33(0.43) |
| Parameter | Unit | Patient 7 | Patient 8 |
| $\log_{10}(\lambda)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | 1.79(0.31) | 2.83(0.11) |
| $\log_{10}(d)$ | $\log_{10}(\frac{1}{day})$ | -1.33(0.31) | -0.44(0.11) |
| $\log_{10}(\beta)$ | $\log_{10}(\frac{mL}{copies \times day})$ | -5.54(0.18) | -5.36(0.12) |
| $\log_{10}(a)$ | $\log_{10}(\frac{1}{day})$ | -0.18(0.15) | 0.42(0.14) |
| $\log_{10}(\gamma)$ | $\log_{10}(\frac{copies \times \mu L}{cells \times mL \times day})$ | 3.66(0.16) | 3.91(0.21) |
| $\log_{10}(\eta)$ | $--$ | -0.17(0.42) | -0.22(0.36) |
| $\log_{10}(\lambda_y)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | -3.46(0.36) | -3.90(0.31) |
| Parameter | Unit | Patient 10 | Patient 11 |
| $\log_{10}(\lambda)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | 2.21(0.37) | 1.65(0.10) |
| $\log_{10}(d)$ | $\log_{10}(\frac{1}{day})$ | -0.96(0.37) | -1.35(0.10) |
| $\log_{10}(\beta)$ | $\log_{10}(\frac{mL}{copies \times day})$ | -5.78(0.17) | -5.54(0.04) |
| $\log_{10}(a)$ | $\log_{10}(\frac{1}{day})$ | 0.00(0.21) | -0.74(0.07) |
| $\log_{10}(\gamma)$ | $\log_{10}(\frac{copies \times \mu L}{cells \times mL \times day})$ | 4.00(0.17) | 3.39(0.01) |
| $\log_{10}(\eta)$ | $--$ | -0.11(0.34) | -0.13(0.06) |
| $\log_{10}(\lambda_y)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | -3.00(0.24) | -4.65(0.09) |

**Table 3.3:** The result comparisons between this thesis and those from literature

| | | This thesis<br>Average MLE<br>(Interpatient Range) | Huang et al [20]<br>Posterior mean<br>(95% CI) | Putter et al [39]<br>Posterior median<br>(Interpatient Range) |
|---|---|---|---|---|
| $\log_{10}(\lambda)$ | $\log_{10}(\frac{cells}{\mu L \times day})$ | 2.47<br>(1.54, 2.88) | 1.97<br>(1.93,2.05) | 0.11<br>(-0.24,0.21) |
| $\log_{10}(d)$ | $\log_{10}(\frac{1}{day})$ | -0.74<br>(-1.35,0.34) | -0.96<br>( -1.01,-0.40) | -2<br>(NA) |
| $\log_{10}(\beta)$ | $\log_{10}(\frac{mL}{copies \times day})$ | -5.41<br>(-5.78,-5.23) | -4<br>( -4.00,-3.15) | -6.80<br>(-6.94,-6.53) |
| $\log_{10}(a)$ | $\log_{10}(\frac{1}{day})$ | 0.01<br>(-0.76,0.42) | -0.42<br>(-0.49,-0.22) | -0.50<br>(-0.62,0.26) |
| $\log_{10}(\gamma)$ | $\log_{10}(\frac{copies \times \mu L}{cells \times mL \times day})$ | 3.77<br>(3.39,4.00) | 2.51<br>(NA) | 4.87<br>(4.83,4.89) |
| $\eta$ | $--$ | 0.73<br>(0.60,0.89) | NA<br>(NA) | 0.71<br>(0.63,0.84) |
| $R_{0pre}$ | NA | 1.93<br>(1.16,3.68) | NA<br>(NA) | 3.90<br>(1.72,4.86) |
| $R_{0post}$ | $--$ | 0.56<br>(0.13,0.76) | NA<br>(NA) | 0.34<br>(0.05,0.60) |
| # of patients | | 10 | 42 | 12 |
| # of measurements per patient | | 98 | 9 | 5 |

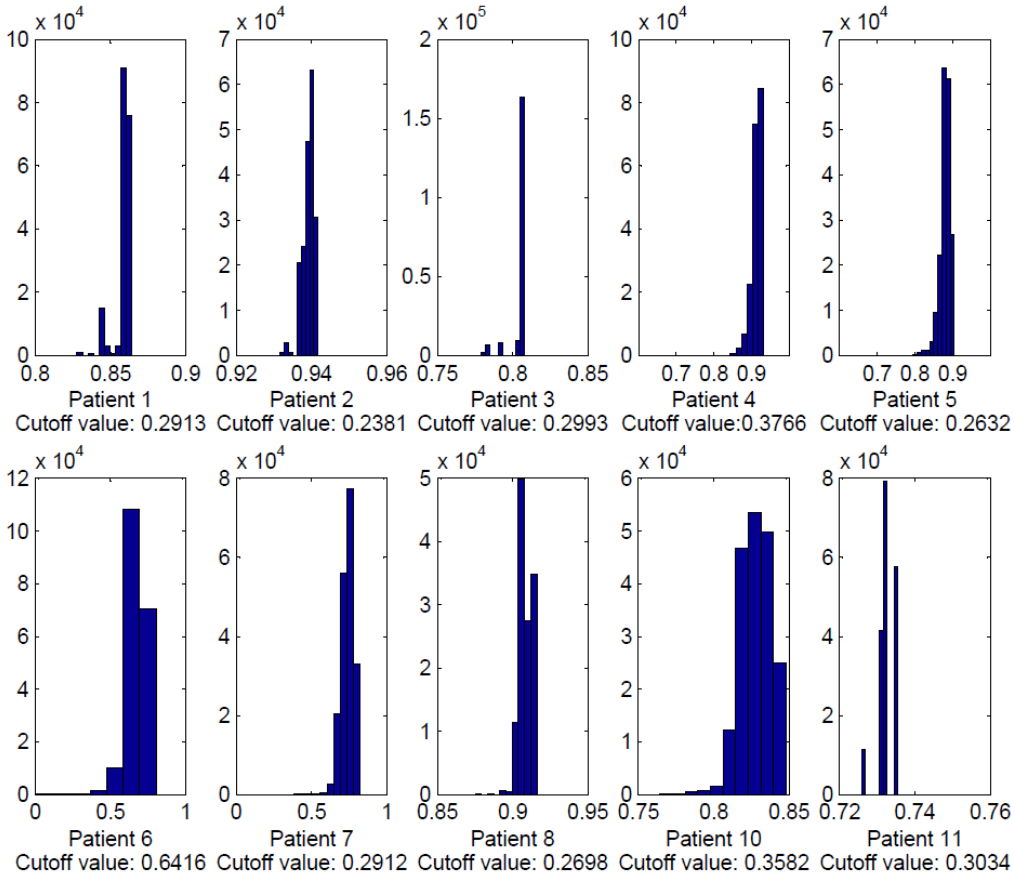


**Figure 3.4: The histogram of coefficient of determination** $R^2$ Cutoff values: The value of $R^2$ corresponding to $P = 0.05$.

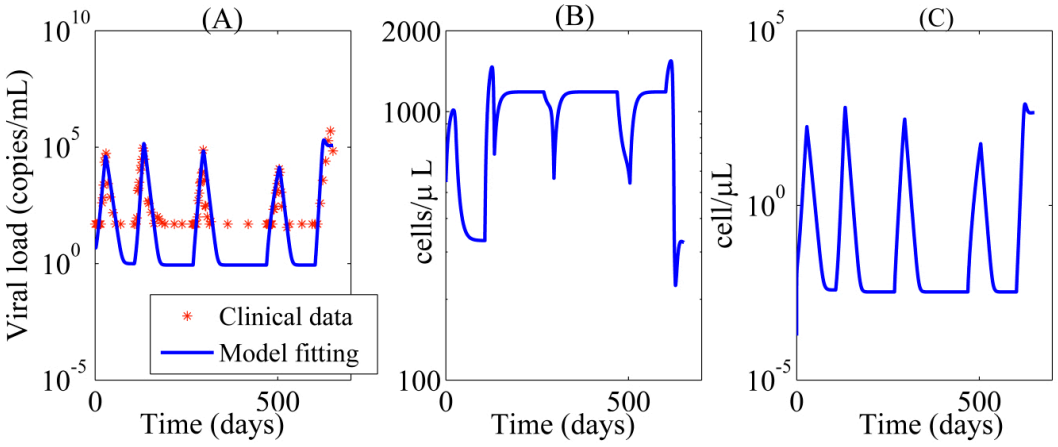


**Figure 3.5: Typical Bayesian Posterior Mean Results:**(A)The fitted viral load curve and viral load data for patient 1; (B)Target cell simulation using fitted parameters; (C)Infected cell simulation using fitted parameters

# Chapter 4

# CONCLUSION AND DISCUSSION

As shown in this study, the parameters in the HIV dynamics model are heavily correlated. The strong correlation between parameters in this model mean that only the distribution of complete parameter sets (as opposed to independent distributions of each parameter) can be considered to accurately represent the fit of the model to the data. The quality and extent of the data available in this study was considerablely higher than any previously published parameter estimation study, allowing for the accurate estimation of the model parameters without using the simplifying assumptions necessitated by the sparsity of data in previous studies. Furthermore, the level of excitation of the dynamics provided by the multiple interruption schedules allowed us to directly identify two parameters not previously identifiable from patient data. The results agree for the most part with previously published data, but the results are more reliable and significant given the dramatic increase in the amount of data available for identification.

This thesis presents the posterior distribution of parameters for a commonly used HIV infection model identified against measured patient data. Analysis of this distribution shows good representation of the data. As shown in the histograms of Fig. 3.4, estimates for all patients except some of patient 6 would be considered statistically significant by the standard F Test ($P < 0.05$), implying that the reported posterior distribution describes the range of feasible parameter values based on the measured data well. Inspection of the data from patient 6 show the data available above the limit of detection are limited compared to the other patients, which can explain the high value of $R^2$ corresponding to a F-test P value of 0.05 for that patient. However, even for this patient, more than half of the estimates are still statistically significant,

and there were 39 measurement points above the limit of detection, significantly more measurements than were available for any previously published parameter estimation study.

The multiple interruptions in the patient data provided the opportunity to quantify the contribution rate of viral reservoirs to the active infected cell compartment. These rates varied widely from patient to patient. The rate was characterized in terms of number of productively infected cells produced per day, and equally well describes such diverse potential reservoir processes as low-level persistent replication, viral blipping, and spontaneous reactivation of quiescent cells. The overall fit is highly sensitive to the rebound time, and the rebound time is uniquely determined by the reservoir contribution rate. It is surprising, therefore, that a single value for the reservoir contribution rate was able to describe the data well over multiple interruptions in the same patient; this strongly suggests that the underlying process represented by the parameter $\lambda_y$ is continuous rather than bursting in nature. The quantification and understanding of the viral reservoir dynamics is of critical importance to understanding the nature of ongoing viral evolution under conditions of effective suppression, and will be a necessary precursor to any attempts to flush the reservoirs and achieve a functional cure for HIV.

In addition to being the most accurate way of describing the fit of a model to data with high levels of measurement uncertainty, the publication of complete parameter distributions identified from patient data also has significant practical importance. A growing number of model-based interventions using variations of the model described in Equation 1.1 are being proposed, including our own methods designed to minimize the risk of resistance emerging during antiviral regimen switching [8,23,24,27,29]. Most of these methods have used either nominal parameters or a single parameter set. The parameter distributions published in this work provide a parameter set against which the robustness of a proposed model-based method to expected patient variation can be tested. The data used in this thesis came from a cohort restricted to patients with good immunological control of the virus under antiviral suppression, so the distribution

of parameters can only be said to be representative of such a subgroup of HIV-infected persons. However, the publication of the identification methods will likely lead to the publication of parameter distributions from patients in other studies in the future, leading to a growing library of virtual patients.

## BIBLIOGRAPHY


[1] Joel N Blankson, Deborah Persaud, and Robert F Siliciano. The challenge of viral reservoirs in HIV-1 infection. *Annu Rev Med*, 53:557–593, 2002.

[2] D. M. Bortz and P. W. Nelson. Model selection and mixed-effects modeling of HIV infection dynamics. *Bull Math Biol*, 68(8):2005–2025, Nov 2006.

[3] T. W. Chun, L. Carruth, D. Finzi, X. Shen, J. A. DiGiuseppe, H. Taylor, M. Hermankova, K. Chadwick, J. Margolick, T. C. Quinn, Y. H. Kuo, R. Brookmeyer, M. A. Zeiger, P. Barditch-Crovo, and R. F. Siliciano. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, 387(6629):183–188, May 1997.

[4] T. W. Chun, D. Finzi, J. Margolick, K. Chadwick, D. Schwartz, and R. F. Siliciano. In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nat Med*, 1(12):1284–1290, Dec 1995.

[5] T. W. Chun, L. Stuyver, S. B. Mizell, L. A. Ehler, J. A. Mican, M. Baseler, A. L. Lloyd, M. A. Nowak, and A. S. Fauci. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad Sci U S A*, 94(24):13193–13197, Nov 1997.

[6] G. Dornadula, H. Zhang, B. VanUitert, J. Stern, L. Livornese, M. J. Ingerman, J. Witek, R. J. Kedanis, J. Natkin, J. DeSimone, and R. J. Pomerantz. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *JAMA*, 282(17):1627–1632, Nov 1999.

[7] G. Dornadula, H. Zhang, B. VanUitert, J. Stern, L. Livornese, M. J. Ingerman, J. Witek, R. J. Kedanis, J. Natkin, J. DeSimone, and R. J. Pomerantz. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *JAMA*, 282(17):1627–1632, Nov 1999.

[8] Jorge Ferreira, Esteban A. Hernandez-Vargas, and Richard H. Middleton. Computer simulation of structured treatment interruption for HIV infection. *Comput Methods Programs Biomed*, 104(2):50–61, Nov 2011.

[9] D. Finzi, M. Hermankova, T. Pierson, L. M. Carruth, C. Buck, R. E. Chaisson, T. C. Quinn, K. Chadwick, J. Margolick, R. Brookmeyer, J. Gallant, M. Markowitz, D. D. Ho, D. D. Richman, and R. F. Siliciano. Identification of

a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*, 278(5341):1295–1300, Nov 1997.

[10] Centers for Disease Control and Prevention (CDC). HIV prevalence estimates–united states, 2006. *MMWR Morb Mortal Wkly Rep*, 57(39):1073–6, Oct 2008.

[11] H. F. Gnthard, D. V. Havlir, S. Fiscus, Z. Q. Zhang, J. Eron, J. Mellors, R. Gulick, S. D. Frost, A. J. Brown, W. Schleif, F. Valentine, L. Jonas, A. Meibohm, C. C. Ignacio, R. Isaacs, R. Gamagami, E. Emini, A. Haase, D. D. Richman, and J. K. Wong. Residual human immunodeficiency virus (HIV) type 1 RNA and DNA in lymph nodes and HIV RNA in genital secretions and in cerebrospinal fluid after suppression of viremia for 2 years. *J Infect Dis*, 183(9):1318–1327, May 2001.

[12] CM Gray, J Lawrence, EA Ranheim, M Vierra, M Zupancic, M Winters, et al. Highly active antiretroviral therapy results in HIV type 1 suppression in lymph nodes, increased pools of naive T cells, decreased pools of activated t cells, and diminished frequencies of peripheral activated HIV type 1-specific CD8+ T cells. *AIDS Res Hum Retroviruses*, 16(14):1357–69, Sep 2000.

[13] R. M. Gulick, J. W. Mellors, D. Havlir, J. J. Eron, C. Gonzalez, D. McMahon, D. D. Richman, F. T. Valentine, L. Jonas, A. Meibohm, E. A. Emini, and J. A. Chodakewitz. Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med*, 337(11):734–739, Sep 1997.

[14] SM Hammer, JJ Eron, P Reiss, RT Schooley, MA Thompson, S Walmsley, et al. Antiretroviral treatment of adult HIV infection: 2008 recommendations of the international AIDS society-usa panel. *JAMA*, 300(5):555–70, Aug 2008.

[15] D. D. Ho. HIV-1 dynamics in vivo. *J Biol Regul Homeost Agents*, 9(3):76–77, 1995.

[16] DD Ho, AU Neumann, AS Perelson, W Chen, JM Leonard, and M Markowitz. Rapid turnover of plasma virions and cd4 lymphocytes in HIV-1 infection. *Nature*, 373(6510):123–6, Jan 1995.

[17] Yangxin Huang. Long-term HIV dynamic models incorporating drug adherence and resistance to treatment for prediction of virological responses. *Computational Statistics & Data Analysis*, 52(7):3765 – 3778, 2008.

[18] Yangxin Huang, Dacheng Liu, and Hulin Wu. Hierarchical bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics*, 62(2):413–423, Jun 2006.

[19] Yangxin Huang, Dacheng Liu, and Hulin Wu. Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics*, 62(2):413–423, June 2006.

[20] Yangxin Huang, Hulin Wu, and Edward P. Acosta. Hierarchical bayesian inference for HIV dynamic differential equation models incorporating multiple treatment factors. *Biom J*, 52(4):470–486, Aug 2010.

[21] Hua Liang, Hongyu Miao, and Hulin Wu. Estimation of constant and time-varying dynamic parameters of HIV infection in a nonlinear differential equation model. *Ann Appl Stat*, 4(1):460–483, Mar 2010.

[22] L. Ljung and S.T. Glad. On global identifiability for arbitrary model parameterizations. *Automatica*, 30(2):265?76, 1994.

[23] R. Luo, L. Cannon, J. Hernandez, M.J. Piovoso, and R. Zurakowski. Controlling the evolution of resistance. *J Process Control*, 21(3):367–378, doi:10.1016/j.jprocont.2010.11.010, 2010.

[24] R Luo, M.J. Piovoso, and R Zurakowski. A generalized multi-strain model of HIV evolution with implications for drug-resistance management. In *Proc. American Control Conference*, pages 2295–2300, 2009.

[25] Rutao Luo, Michael J Piovoso, Javier Martinez-Picado, and Ryan Zurakowski. Optimal Antiviral Switching to Minimize Resistance Risk in HIV Therapy. *PLoS ONE*, 6(11):e27047, 2011.

[26] Rutao Luo, Michael J. Piovoso, Javier Martinez-Picado, and Ryan Zurakowski. Hiv model parameter estimates from interruption trial data including drug efficacy and reservoir dynamics. *PloS One*, 7(7):e40198, 2012.

[27] Rutao Luo and R. Zurakowski. A new strategy to decrease risk of resistance emerging during therapy switching in HIV treatment. In *Proc. American Control Conference*, pages 2112–2117, 11–13 June 2008.

[28] Martin Markowitz, Michael Louie, Arlene Hurley, Eugene Sun, Michele Di Mascio, Alan S Perelson, and David D Ho. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. *J Virol*, 77(8):5037–5038, April 2003.

[29] Jorge L. Martinez-Cajas and Mark A. Wainberg. Antiretroviral therapy : optimal sequencing of therapy to avoid resistance. *Drugs*, 68(1):43–72, 2008.

[30] J Martinez-Picado, A V Savara, L Sutton, and R T D'Aquila. Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. *Journal of virology*, 73(5):3744–52, May 1999.

[31] J Martinez-Picado, AV Savara, L Shi, L Sutton, and RT D'Aquila. Fitness of human immunodeficiency virus type 1 protease inhibitor-selected single mutants. *Virology*, 275(2):318–22, Sep 2000.

[32] Hongyu Miao, Carrie Dykes, Lisa M Demeter, and Hulin Wu. Differential equation modeling of HIV viral fitness experiments: model identification, model selection, and multimodel inference. *Biometrics*, 65(1):292–300, Mar 2009.

[33] G. W. Nelson and A. S. Perelson. Modeling defective interfering virus therapy for AIDS: conditions for div survival. *Math Biosci*, 125(2):127–153, Feb 1995.

[34] Sarah Palmer, Frank Maldarelli, Ann Wiegand, Barry Bernstein, George J Hanna, Scott C Brun, Dale J Kempf, John W Mellors, John M Coffin, and Martin S King. Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. *Proc Natl Acad Sci U S A*, 105(10):3879–3884, Mar 2008.

[35] Sarah Palmer, Ann P Wiegand, Frank Maldarelli, Holly Bazmi, JoAnn M Mican, Michael Polis, Robin L Dewar, Angeline Planta, Shuying Liu, Julia A Metcalf, John W Mellors, and John M Coffin. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol*, 41(10):4531–4536, Oct 2003.

[36] A Perelson. Dynamics of HIV infection of CD4+ T cells. *Mathematical Biosciences*, 114(1):81–125, March 1993.

[37] A. S. Perelson, P. Essunger, Y. Cao, M. Vesanen, A. Hurley, K. Saksela, M. Markowitz, and D. D. Ho. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*, 387(6629):188–191, May 1997.

[38] A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271(5255):1582–1586, Mar 1996.

[39] H Putter, S H Heisterkamp, J M A Lange, and F de Wolf. A Bayesian approach to parameter estimation in HIV dynamical models. *Stat Med*, 21(15):2199–2214, August 2002.

[40] B Ramratnam, S Bonhoeffer, J Binley, A Hurley, L Zhang, J E Mittler, M Markowitz, J P Moore, A S Perelson, and D D Ho. Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *The Lancet*, 354(9192):1782–1785, November 1999.

[41] RM Ribeiro and S Bonhoeffer. Production of resistant HIV mutants during antiretroviral therapy. *Proc Natl Acad Sci USA*, 97(14):7681–6, Jul 2000.

[42] RM Ribeiro, H Mohri, DD Ho, and AS Perelson. In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not CD8+ T cells depleted? *Proc Natl Acad Sci U S A*, 99(24):15572–7, Nov 2002.

[43] J. F. Ritt. *Differential algebra*. American Mathematical Society, 1950.

[44] Libin Rong and Alan S. Perelson. Modeling HIV persistence, the latent reservoir, and viral blips. *Journal of Theoretical Biology*, 260(2):308 – 331, 2009.

[45] Libin Rong and Alan S Perelson. Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLoS Comput Biol*, 5(10):e1000533, Oct 2009.

[46] L. Ruiz, G. Carcelain, J. Martnez-Picado, S. Frost, S. Marfil, R. Paredes, J. Romeu, E. Ferrer, K. Morales-Lopetegi, B. Autran, and B. Clotet. HIV dynamics and t-cell immunity after three structured treatment interruptions in chronic HIV-1 infection. *AIDS*, 15(9):F19–F27, Jun 2001.

[47] Maria Pia Saccomani, Stefania Audoly, Giuseppina Bellu, and Leontina D'Angio. Examples of testing global identifiability of biological and biomedical models with the daisy software. *Computers in Biology and Medicine*, 40(4):402 – 407, 2010.

[48] Maria Pia Saccomani, Stefania Audoly, and Leontina D'Angio. Parameter identifiability of nonlinear systems: the role of initial conditions. *Automatica*, 39(4):619 – 632, 2003.

[49] L. K. Schrager and M. P. D'Souza. Cellular and anatomical reservoirs of HIV-1 in patients receiving potent antiretroviral combination therapy. *JAMA*, 280(1):67–71, Jul 1998.

[50] Janet D Siliciano, Joleen Kajdas, Diana Finzi, Thomas C Quinn, Karen Chadwick, Joseph B Margolick, Colin Kovacs, Stephen J Gange, and Robert F Siliciano. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med*, 9(6):727–728, Jun 2003.

[51] Davey M Smith, Joseph K Wong, Hai Shao, George K Hightower, Stephanie H T Mai, Joseph M Moreno, Caroline C Ignacio, Simon D W Frost, Douglas D Richman, and Susan J Little. Long-term persistence of transmitted HIV drug resistance in male genital tract secretions: implications for secondary transmission. *J Infect Dis*, 196(3):356–360, Aug 2007.

[52] M. A. Stafford, L. Corey, Y. Cao, E. S. Daar, D. D. Ho, and A. S. Perelson. Modeling plasma virus concentration during primary HIV infection. *J Theor Biol*, 203(3):285–301, Apr 2000.

[53] M Trignetti, T Sing, V Svicher, M Mercedes Santoro, F Forbici, R D'arrigo, et al. Dynamics of nrti resistance mutations during therapy interruption. *AIDS research and human retroviruses*, 25(1):57–64, Jan 2009.

[54] X. Wei, S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, and B. H. Hahn. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*, 373(6510):117–122, Jan 1995.

[55] Hulin Wu, Yangxin Huang, Edward P Acosta, Susan L Rosenkranz, Daniel R Kuritzkes, Joseph J Eron, Alan S Perelson, and John G Gerber. Modeling long-term HIV dynamics and antiretroviral response: effects of drug potency, pharmacokinetics, adherence, and drug resistance. *J Acquir Immune Defic Syndr*, 39(3):272–283, Jul 2005.

[56] Hulin Wu, Haihong Zhu, Hongyu Miao, and Alan S Perelson. Parameter identifiability and estimation of HIV/AIDS dynamic models. *Bull Math Biol*, 70(3):785–799, Apr 2008.

[57] X. Xia. Estimation of HIV/AIDS parameters. *Automatica*, 39(11):1983–1988, 2003.

[58] X. Xia and C. H. Moog. Identifiability of nonlinear systems with application to HIV/AIDS models. *IEEE T Automat Contr*, 48(2):330–336, 2003.