

Submit to *Biometrics*
(Consultant's Forum)

**A Comparison Study of Models and Fitting
Procedures for Biphasic Viral Dynamics in
HIV-1 Infected Patients Treated with
Antiviral Therapies**

A. Adam Ding and Hulin Wu

June, 1998

Corresponding Author:

Hulin Wu, Ph.D.

Frontier Science & Technology

Research Foundation, Inc.

1244 Boylston Street, Suite 303

Chestnut Hill, MA 02467-2104

Phone: (617) 632-5738

Fax: (617) 632-2001

Email: wu@sdac.harvard.edu

A Comparison Study of Models and Fitting Procedures for Biphase Viral Dynamics in HIV-1 Infected Patients Treated with Antiviral Therapies

A. Adam Ding

Department of Mathematics, Northeastern University, Boston, Massachusetts 02115,
U.S.A.

and

Hulin Wu*

Statistical and Data Analysis Center, Harvard School of Public Health, Frontier Science
& Technology Research Foundation, Inc., 1244 Boylston Street, Suite 303, Chestnut
Hill, Massachusetts 02467, U.S.A.

SUMMARY

The study of HIV dynamics is one of the most important developments in recent AIDS research. It has led to a new understanding of the pathogenesis of HIV infection. But although important findings in HIV dynamics have been published in prestigious scientific journals such as *Science* and *Nature* in the last three years (Ho et al., 1995; Wei et al., 1995; Perelson et al., 1996 and 1997), the model-fitting procedures used in these publications have not been studied in any detail. In this paper we evaluate the performance of four model-fitting procedures proposed and used in biphase HIV dynamic data analysis via extensive Monte Carlo simulations. We propose some guidelines for practitioners to select an appropriate method for their own data analysis. Real data examples from an AIDS clinical trial are provided as illustrations.

* Corresponding author's email address: wu@sdac.harvard.edu

Key Words: AIDS modeling; HIV infection; Monte Carlo simulation; Nonlinear regression; Viral load.

1. Introduction

Recently there has been a great interest in estimating HIV dynamic parameters in order to acquire a greater understanding of the pathogenesis of HIV infection. HIV dynamic models can also provide theoretical principles to guide the development of treatment strategies for HIV-infected patients (Ho et al., 1995; Wei et al., 1995; Perelson et al., 1996 and 1997; Wu et al., 1997, 1998b). Essunger et al. (1997) and Ding and Wu (1998) have also proposed using viral dynamics to evaluate the efficacy of anti-HIV therapies.

Although important findings in HIV dynamics have been published in prestigious scientific journals such as *Science* and *Nature* in the last three years (Ho et al., 1995; Wei et al., 1995; Perelson et al., 1996 and 1997), the model-fitting procedures used in these publications have not been studied in any detail. In case where clinical data were not sufficient for estimation of all parameters, an ad hoc substitution method (substituting unknown parameters from other studies) was used (Perelson et al., 1997). Wu and Ding (1998) proposed another procedure to reduce complicated HIV dynamic models to estimable functions by reparameterization, while Perelson et al. (1996, 1997) proposed using a pre-treatment steady-state assumption to reduce the number of estimated parameters. However, the steady-state condition is difficult to validate and may not hold for some situations, such as newborn infants (Luzuriaga et al., 1998; Wu et al. 1998c). Thus it becomes more and more important to study these proposed model-fitting procedures and to select an appropriate and efficient procedure for different cases in order to avoid misleading results in this critical field of AIDS research.

In this paper we evaluate the proposed model-fitting procedures and model assumptions using extensive Monte Carlo simulations. Our purpose is to provide guidelines for practitioners to select appropriate HIV dynamic model-fitting procedures based on the situations in their own studies. In Section 2, we introduce the four major biphasic HIV dynamic models and parameter estimation procedures used and proposed in the literature (Perelson et al. 1997; Wu and Ding, 1998; Luzuriaga et al., 1998; Melvin et al., 1998). In Section 3, we compare the four model-fitting procedures under three viral load mea-

surement schedules. The performance of the four procedures is evaluated based on these simulation results. Real data examples from an AIDS clinical trial are used to illustrate the procedure selection principles in Section 4. We conclude the paper with some discussions.

2. HIV Dynamic Models and Estimation Procedures

The use of a system of differential equations to describe the interaction between HIV and its host cells can be traced back to the end of the 1980s (Merrill, 1987; Anderson and May, 1989; Perelson, 1989). A deterministic model for HIV dynamics after initiation of antiviral treatments can be written as (Perelson, Kirschner and Boer, 1993; Kirschner and Perelson, 1995; Wu and Ding, 1998; Ding and Wu, 1998):

$$\begin{aligned}
\frac{d}{dt}T_l &= (1 - \gamma_l)k_lTV_I - \delta_lT_l, \\
\frac{d}{dt}T_p &= (1 - \gamma_p)k_pTV_I - \delta_pT_p, \\
\frac{d}{dt}V_I &= (1 - \eta_0)((1 - \eta_l)N_l\delta_lT_l + (1 - \eta_p)N_p\delta_pT_p) - cV_I, \\
\frac{d}{dt}V_{NI} &= [(\eta_0 + \eta_l(1 - \eta_0))N_l\delta_lT_l + (\eta_0 + \eta_p(1 - \eta_0))N_p\delta_pT_p] - cV_{NI}.
\end{aligned} \tag{1}$$

where $T, T_l, T_p, V_I,$ and V_{NI} denote the concentration of uninfected target cells, long-lived infected cells, productively infected cells, infectious virus and noninfectious virus, respectively. The target cells (T) can be infected and become long-lived infected cells (T_l) at a rate of k_lV_I without treatment and $(1 - \gamma_l)k_lV_I$ during treatment with reverse transcriptase inhibitor (RTI) drugs. The target cells can also become productively infected cells (T_p) at a rate of k_pV_I without treatment and $(1 - \gamma_p)k_pV_I$ during treatment with RTI drugs. Parameters γ_l and γ_p are the treatment effects of the RTI drugs ($\gamma_l = \gamma_p = 1$ means perfect treatment and $\gamma_l = \gamma_p = 0$ means no treatment effect). We denote $1 - \eta_0$ as the proportion of infectious virus produced by infected cells without the intervention of protease inhibitor (PI) antiviral drugs. It will be reduced by factors of $1 - \eta_l$ and $1 - \eta_p$ for long-lived and productively infected cells, respectively, during treatment with PI drugs. Parameters η_l and η_p are the treatment effects of the PI drugs, and N_l and N_p are the average numbers of virions produced per long-lived or productively infected

cell, respectively, during their lifetimes. Parameters $\mu_T, \delta_l, \delta_p$ and c are the death rates of T, T_l, T_p and virus, respectively. More details on these notations and assumptions can be found in Perelson, Kirschner and Boer (1993), Kirschner and Perelson (1995), Wu and Ding (1998) and Ding and Wu (1998). This model will be used to generate the true trajectory of viral decay in our simulations (Section 3).

This model had not been used in clinical trials to estimate HIV dynamic parameters, due to the complexity of the model, until the recent development of simplification and approximation techniques (Ho et al., 1995; Wei et al., 1995; Perelson et al., 1996 and 1997; Wu and Ding, 1998). Only two infected cell compartments (productively infected cells and long-lived infected cells) can be identified based on plasma viral load data due to the limitations of current assays. These two compartments are believed to produce a biphasic viral decay during treatment with potent antiviral therapies (Perelson et al., 1997). Under the assumption of a constant target cell concentration, a solution for the total viral load can be obtained as (Ding and Wu, 1998),

$$V(t) = P_0 e^{-d_0(t-t_d)} + P_1 e^{-d_1(t-t_d)} + P_2 e^{-d_2(t-t_d)}, \quad t \geq t_d \quad (2)$$

where t is treatment time and t_d denotes the time of intracellular and pharmacological delay (Perelson et al., 1996; Herz, 1996). Under the assumption of perfect therapy, the exponential decay rates, d_0, d_1 , and d_2 are the exact death rates of free HIV virions, productively infected cells and long-lived infected cells, c, δ_p , and δ_l , respectively (Perelson et al., 1997; Wu and Ding, 1998). If therapy is not perfect, approximate formulas for d_0, d_1 , and d_2 are also derived by Ding and Wu (1998). Parameters P_0, P_1 , and P_2 are coefficient parameters for each corresponding compartment. In practice, the target cell concentration may not remain constant during long-term clinical trials and should be modeled by a complicated differential equation, but equation (2) is still a good approximate solution (Wu and Ding 1998).

In model (2), there are 7 unknown parameters, P_i and $d_i, i = 0, 1, 2$, and t_d . However, not all of these parameters can be identified based on repeated measurements of viral load $V(t), t = 1, 2, \dots, n$. The time of intracellular and pharmacological delay, t_d , is almost

impossible to estimate accurately unless we have frequent measurements of both viral load and pharmacokinetics/pharmacodynamics (Herz et al. 1996), which are not practical due to limitations on the amount of blood drawn within a fixed time period. Since most clinical trials are not designed to have intensive measurements of viral load during the first 1 or 2 days, as in Perelson et al. (1996), parameter d_0 ($\approx c$, the clearance rate of free virions) also cannot be obtained for each patient. In order to estimate other parameters such as d_1 and d_2 based on biphasic plasma viral decay data, several ad hoc methods have been proposed and used (Perelson et al., 1997; Wu and Ding, 1998).

Method 1 (NAIVE): Since the term $P_0e^{-d_0(t-t_d)}$ in model (2) is negligible compared to the other two terms after 1 or 2 days of treatment and t_d is as short as only several hours (Perelson et al., 1996), we can simply ignore them and fit all the data to a biphasic exponential model,

$$V(t) = P_1e^{-d_1t} + P_2e^{-d_2t}. \quad (3)$$

Due to its simplicity, we refer to this method as a “naive” method. It was used to fit the data from a study of HIV-1 infected infants and children by Melvin et al. (1998). This method certainly would introduce bias into the estimates of parameters P_1, P_2, d_1 and d_2 . However, since all the data are used and the model is parsimonious, the variance of the estimates may be small. The trade-off between the bias and the variance of the estimates has to be considered carefully for this method.

Method 2 (WD): In order to reduce the bias in the above naive method, Wu and Ding (1998) proposed ignoring the data obtained during the first 1 or 2 days of treatment and fitting the bi-exponential model (3) using only the data after the small shoulder produced by the term $P_0e^{-d_0(t-t_d)}$ and the pharmacological delay, t_d . Since we ignore the data that are not on the bi-exponential model, the estimation bias will be reduced. However, since the data are not sufficiently used, the variance of the estimates will be larger. Thus we still need to consider the trade-off between the bias and variance. We refer to this method

as the Wu-Ding method (WD).

Method 3 (PNSS): Perelson et al. (1997) proposed and used a substitution method, i.e., they substituted the unknown parameters d_0 ($\approx c$) and t_d by estimates from a previous study. In Perelson et al. (1996), the estimate of c ranges from 2.06 to 3.81 with a mean of 3.07. The estimate of t_d ranges from 2 to 6 hours with a mean of 3.6 hours. Thus Perelson et al. (1997) substituted $d_0 \approx c$ by its mean 3.0. Because t_d turns out to have very little effect on the estimates of d_1 and d_2 , it is simply substituted by $t_d = 0$. Thus only 5 parameters in model (2) need to be estimated. Nevertheless, this mean substitution method may still introduce bias into the estimates of parameters, since the parameters c and t_d for a particular patient may depart far from the mean estimates of the previous study. This method uses all the data, but one more parameter needs to be estimated compared to Method 1 and 2. Since the steady-state assumption is not used in this method, we refer to this method as the Perelson non-steady-state method (PNSS). Luzuriaga et al. (1998) used this method in their study.

Method 4 (PSS): To reduce the number of parameters that need to be estimated, Perelson et al. (1997) assumed a steady-state before treatment. Then model (2) can be re-written as

$$V(t) = V_B[Ae^{-d_1(t-t_d)} + Be^{-d_2(t-t_d)} + (1 - A - B)e^{-d_0(t-t_d)}]. \quad (4)$$

where $d_0 = c$, $d_1 = \delta_p$, $d_2 = \delta_l$ under the assumption of perfect therapy, and $A = \alpha/(c - \delta_p)$ and $B = (c - \alpha)/(c - \delta_l)$ (Perelson et al., 1997; Wu and Ding, 1998). Here $\alpha = NkT_B$, where N is the number of virions produced for each infected cell in its lifetime, k is the infection rate, and V_B and T_B are the baseline concentrations of virions and target cells. Again we can substitute $d_0 = c = 3$ and $t_d = 0$. Now only 4 parameters, V_B , α , d_1 and d_2 need to be estimated. However, this method has the same problem as Method 3, and in addition, the steady-state assumption is difficult to validate in practice. We refer to this method as the Perelson steady-state method (PSS).

Intuitively, the NAIVE method may give the smallest variance for the estimates since all data are used and the fewest number of parameters is estimated. Also, due to its simple form of solution, the convergence rate of numerical computation would be the highest for the NAIVE method. However, since it uses the data on the small shoulder during the first one or two days of treatment (Perelson et al., 1996; Herz, 1996; Wu and Ding, 1998) and these data do not follow the biphasic decay model, it would produce the largest bias for the estimates. The WD method may give the smallest bias for the estimates since it ignores the data on the small shoulder where the biphasic model does not apply. The WD method also excludes the possible bias due to the biased substitution of c and t_d and that produced by the steady-state assumption. However, as mentioned previously, this method may produce a larger estimation variance compared to other methods since it does not use all the data. It may also run into numerical instability problems if the data are not sufficient after excluding the data on the small shoulder.

Since the PNSS method needs to estimate one more parameter compared to the other methods, the convergence rate might be lowest. The convergence rate for the PSS method is almost as good as the NAIVE method unless the steady-state assumption departs too far from the truth, since the number of parameters and the number of data points are the same for both the PSS and NAIVE methods. The bias of both the PNSS and PSS methods depends on the bias of the substituted values of c and t_d . The bias of the PSS method also depends on the departure of the steady-state assumption from the truth. It is very difficult to intuitively evaluate the estimate variance from the PNSS and PSS methods due to the complicated trade-off between the number of estimated parameters, the number of data points, and the steady-state assumption.

All four methods have their advantages and drawbacks. The picture is murkier when we evaluate their performance in terms of estimate bias, variance, and numerical stability simultaneously. Therefore a simulation study is necessary to quantify the bias, variance and numerical stability for the four methods in various cases to provide a guideline for practitioners.

3. Comparison Studies

3.1 Factors Affecting the Performance of Estimation Methods

Since different assumptions are made for the different estimation methods introduced in the above section, the performance of these methods may differ under different situations. Therefore we may not be able to find one method which always outperforms the others in all cases. Although many factors may affect the performance of these methods, we consider three major factors in our comparison studies, i.e., the steady-state assumption, the substituted values of c and t_d , and the sampling schedule.

Pre-treatment steady-state was assumed in the PSS method (Perelson et al., 1997), i.e., the interaction between virus and its host cells was assumed to be in a steady-state before initiation of antiviral therapy. This assumption may not be true in cases such as infected adults during acute (primary) infection (Phillips, 1996) and vertically infected infants and children. The PSS method may only be favored if the pre-treatment condition is truly in a steady-state. We can evaluate the steady-state assumption using the production/clearance ratio, R . If the steady-state holds, $R = 1$, i.e., the viral production equals the clearance and the patient's viral load should be stable. If $R > 1$, the viral production is greater than the clearance; the patient's viral load is increasing. On the other hand, if $R < 1$, the viral production is less than the clearance; the patient's viral load is decreasing. Thus, pre-treatment viral load measurements may be used to determine whether the steady-state assumption holds.

In the PNSS and PSS methods, the parameter c is substituted by an estimate from a previous study (usually $c = 3$ from Perelson et al., 1996). The substituted value of c may affect the performance of these methods. The true c for different patients and in different studies may differ from the substituted value (it is more likely greater than 3 according to a new study by Zhang et al., 1998). The substituted value of t_d may also affect the performance of the PNSS and PSS methods, but the effect of t_d on the estimates of d_1 and d_2 is small.

The sampling schedule of viral load measurements may affect the convergence and the estimates of the four methods. If the sampling schedule is sparse, numerical convergence is a serious problem. In this case, the NAIVE and PSS methods may be favored since the smallest number of parameters needs to be estimated and all the data points are used. On the other hand, if the sampling schedule is frequent enough, the numerical convergence may no longer be a problem. The trade-off between the variance and bias of the estimates needs to be considered carefully in this case.

3.2 Design of Simulation Experiments

To evaluate the performance of the four methods, we conducted Monte Carlo simulations for various cases. First, we generated the true trajectory of total viral load $V(t) = V_I(t) + V_{NI}(t)$ along time t based on model (1), with parameters obtained from the published literature (Table 1) or from the estimates of real data (see next section). Second, we generated the observed viral load by adding measurement error to the true value of $V(t)$. In medical research, people prefer to use the \log_{10} scale for viral load. After log transformation, the measurement error approximately follows a normal distribution with a constant variance. Thus we generated the observation of viral load at measurement times t_1, t_2, \dots, t_n from

$$y_i = \log_{10}[V(t_i)] + \varepsilon_i, \quad i = 1, \dots, n \quad (5)$$

where ε_i are i.i.d. measurement errors with $N(0, \sigma^2)$. Based on estimates from ACTG 356 data (Luzuriaga et al., 1998), we chose $\sigma = 0.17$.

To evaluate the effect of the three major factors (mentioned in Subsection 3.1) on the performance of the estimation methods, we simulated the different situations with different combinations of the three factors. For the factor of the steady-state assumption, we selected the steady-state, $R = 1$; the state of viral load decreasing, $R = 0.5$; and the state of viral load increasing, $R = 3$. For the clearance rate of free virions, we substituted $c = 3$ in our model-fitting when using the PNSS and PSS methods as in Perelson et al. (1997). But the true values of c were taken as $c = 2, 3$ and 6 respectively when we

Table 1: Parameter Values Used in the Simulation Study

Parameters	Values	Parameters	Values
δ_p	0.95	k_l	$2 \times 10^{-5} \times R$
δ_l	0.05	k_p	$3.8 \times 10^{-4} \times R$
η_0	0.9	$N_l = N_p$	200
$\eta_l = \eta_p$	0.9	$V_I(0)$	5000
$\gamma_l = \gamma_p$	0.8	$T(0)$	1000

simulated the true viral load data. We assumed that $t_d = 0$ and $t_d = 6$ hours respectively in our model-fitting, but the true value of t_d was generated from a uniform distribution between 2 and 10 hours.

Three sampling schedules were used in our simulation studies. The first schedule is very frequent (ideal), which may not be practical, but is used for comparison purposes. The second and third schedules followed the actual schedules of AIDS Clinical Trial Group protocols 356 and 315 (Luzuriaga et al., 1998; Wu et al., 1998b).

- Schedule 1 (Ideal): Days 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 17, 21, 28.
- Schedule 2 (ACTG 356): Hours 0, 3, 8, and Days 1, 3, 7, 14, 28, 56, 84.
- Schedule 3 (ACTG 315): Days 0, 2, 7, 10, 14, 21, 28, 56.

Based on the above parameter specification and sampling schedule, we simulated the viral load observations y_i , $i = 1, 2, \dots, n$ using models (1) and (5). Then we used the nonlinear least squares method to estimate the viral dynamic parameters based on the four methods introduced in Section 2. We are interested in the estimates of the viral decay parameters, d_1 and d_2 . We compared the estimated d_1 and d_2 from the four methods to the true values (estimated from the data without measurement error). We repeated each simulation case 500 times. The results are summarized in the next subsection.

3.3 Results of Simulation Experiments

The simulation results for $R = 0.5, 1, \text{ and } 3$ are reported in Table 2-4 respectively. Since the results using $t_d = 0$ and $t_d = 6$ are similar, only the results for $t_d = 0$ are reported in these tables. The numerical convergence rates based on the 500 simulation runs for the different methods are given in the fourth column of these tables. We report the bias and standard deviation (STD) of the estimates for d_1 and d_2 as well as the standard error (SE) which is defined as the square root of the mean-squared error (MSE):

$$SE = \sqrt{\text{mean}(\hat{d}_1 - d_1)^2} = \sqrt{\text{Bias}^2 + \text{STD}^2} \quad (6)$$

for one individual, and

$$SE = \sqrt{\text{Bias}^2 + \text{STD}^2/n} \quad (7)$$

for a population sample size n . The results for one individual, $n = 16$ (ACTG 356 sample size), and $n = 50$ (ACTG 315 sample size) are given in the table. To evaluate the performance of the estimation methods, we can compare their biases and standard deviations separately or compare their SEs directly.

In Table 2-4, we use boldface to indicate the smallest bias, standard deviation and SE among the four methods for different situations (if the convergence rate is lower than 85%, the method is not considered). From these results, it is clear that we may not be able to find one method which is consistently the best in terms of convergence rate, bias, standard deviation and SE. However, when we consider these factors together, we suggest one or two favored methods for each case (boldface in Table 2-4).

As we expected, the numerical convergence is reasonably good in most cases, except that the convergence rates of the PNSS and WD methods are low when the sampling schedule is sparse (ACTG 315). The estimates of d_2 from the four methods are quite similar, although the PNSS estimates are a little better than the other three methods when the sampling schedule is frequent, and the PSS estimates are a little better in the other two sampling schedules (ACTG 356 and ACTG 315). These results are what we expected since these four estimation methods differ mainly in how they deal with

the earlier data. This difference has little effect on the estimation of d_2 , which depends mainly on the data during the second phase. Thus the evaluation and selection of the four methods would be based on the estimates of d_1 .

The diversity of the performance of the four methods is great in estimating d_1 . But we still can see a clear trend. The individual estimate from the NAIVE method is the best among the four methods when the viral load is decreasing before treatment ($R < 1$) or when the true clearance rate of free virions, c , is larger than the substituted value in the PNSS and PSS methods in the pre-treatment steady-state condition; otherwise, the PSS method is preferable. For population estimates, the WD method is preferable when the sampling schedule is not too sparse; otherwise, the PSS method is better. However, due to the diverse simulation results for different situations, we suggest that the reader conduct a simulation based on the situation in their study and select an appropriate method for their data analysis using the ideas outlined in this section.

4. Real Data Example

Ideally, the factors which affect the performance of the estimation methods are homogeneous for the patient population in a study. In that case, one method may be selected for the whole population based on the population SE with sample size n in Table 2-4. However, in some cases, differences in these factors exist among the patient population. The best estimation method is different for different individuals. In this case, the individual SE should be used to select the estimation method. Data from ACTG 356 are used to illustrate this idea as follows.

Sixteen infants aged from 15 days to 2 years enrolled in ACTG 356 and initiated combination antiretroviral therapy with zidovudine (ZDV), lamivudine (3TC), and nevirapine (NVP). Plasma viral load measurements (V) were taken just prior to therapy and at 3 and 8 hours; at 1, 3, 7, 14, and 28 days; and then every 28 days. A biphasic viral decay was observed for all patients. More details on clinical results from this study are reported in Luzuriaga et al. (1998). We obtained individual estimates of the viral decay rates, d_1

and d_2 , from all four methods. However, we reached convergence for all 16 patients using the NAIVE method; for 14 patients using the WD method; and for 15 patients using the PNSS and PSS methods, respectively. In order to suggest one estimate for each patient to medical investigators, we need to evaluate the estimates from the four methods based on the simulation results in Table 2-4. However, two major factors, c and R , discussed in previous sections, need to be evaluated patient by patient. Since a new study (Zhang et al., 1998) has shown that the clearance rate of free virions, c , is probably greater than 3 (the value we used in our simulation), we may choose $c \geq 3$.

It is not easy to accurately estimate R . Fortunately there are usually several viral load measurements prior to therapy in most AIDS clinical trials. These measurements may be used to roughly determine R . In the ACTG 356 study, we have 2 or 3 data points during patient screening and at baseline prior to therapy. A linear regression of these viral load measurements, y_i , versus time, t_i , may be used to determine whether the viral load is stable ($R = 1$) before treatment, i.e.,

$$y_i = y_0 + S \times t_i + \varepsilon_i,$$

where y_i is usually a log scale viral load measurement. The slope S can be used to roughly decide R . If $S = 0$, the viral load of the patient is likely to be in a steady-state ($R = 1$); if $S < 0$, the viral load is decreasing ($R < 1$); if $S > 0$, the viral load is increasing ($R > 1$). From ACTG 356, we selected 3 individual patients who represent 3 different cases to discuss in detail. The estimates of d_1 and d_2 from the four methods for these 3 individuals are given in Table 5.

The viral load data for these 3 patients are shown in Figure 1. The estimated slopes for the three subjects are $S = -0.095$, $S = 0.009$, and $S = 0.155$, respectively. The first subject with $S = -0.095$ may be considered as non-steady-state with viral load decreasing ($R < 1$); the second subject with $S = 0.009 (\approx 0)$ may be considered as steady-state ($R = 1$); and the subject with $S = 0.155$ may be considered as non-steady-state with viral load increasing ($R > 1$). Notice that the determination of R in this example is very rough and it is just used for purposes of illustration. However, if the data before treatment

are sufficient, a formal statistical test may be used to determine whether $R < 1$, $R = 1$ or $R > 1$.

For the first subject, we looked up Table 2 for the sampling schedule from ACTG 356. When $c = 3$ or 6, the NAIVE method gave the smallest SE in the estimate of d_1 for one individual ($n = 1$) and almost the same SE in the estimate of d_2 . Thus we propose the estimate from the NAIVE method for this individual. For the second subject, we looked up Table 3, since the viral load of this patient was approximately in a steady-state. The NAIVE method also gave the smallest SE in the estimate of d_1 for one individual ($n = 1$) and almost the same SE in the estimate of d_2 when $c = 3$ or 6. Therefore we prefer the estimate from the NAIVE method for this subject too. For the third subject, we looked up Table 4. We suggest the PNSS method for this subject, since the PNSS method gave the smallest bias and SE for the estimates of d_1 and d_2 when $c = 3$ or 6, although the PSS method produces a slightly better SE when $c = 6$ and $n = 1$. The fitted curves from the suggested methods for these 3 subjects are shown in Figure 1.

Above, we have illustrated how to select the appropriate estimation method for obtaining individual estimates. However, one might like to select one method for the whole population in a study. In this case, the factor R needs to be determined for the whole population and one looks up the appropriate simulation tables (or does one's own simulations) based on the sample size n .

5. Discussion

Several factors and assumptions affect the performance of estimation procedures for biphasic HIV dynamic parameters. The selection of these methods for a particular study should not be arbitrary. To select the most appropriate and efficient method, Monte Carlo simulations may provide a useful tool. If a study is similar to one of the simulation experiments in this paper, the simulation results in Table 2-4 can be used to select an appropriate method. If one has a different study, we suggest conducting a simulation study based on the actual situation and the simulation procedure illustrated in Section 3.

We used nonlinear least squares methods in our comparison study and our comparisons were based on individual or population averages of parameter estimates. The comparison for the population estimate based on hierarchical mixed-effect models (Wu, Ding and DeGruttola, 1998a; Wu and Ding, 1998) may be similar, but further simulation studies may be needed under this setting.

ACKNOWLEDGMENT

This work was supported by NIAID/NIH grants No. R29 AI43220 and No. U01 AI38855. The authors would like to thank the teams of AIDS Clinical Trial Group Protocols 356 and 315 for permission to use their HIV-1 RNA data.

REFERENCES

- Anderson, R.M. and May, R.M. (1989). Complex dynamical behavior in the interaction between HIV and the immune system. In *Cell to Cell Signaling: From Experiments to Theoretical Models*, A. Goldbeter, Ed., Academic, New York.
- Essunger P, Markowitz M, Ho DD, Perelson AS: Efficacy of drug combination and dosing regimen in antiviral therapy. *The First International Workshop on HIV Drug Resistance, Treatment Strategies and Eradication*, Abstract 73.
- Ding, A.A. and Wu, H. (1998). Assessing efficacy of anti-HIV therapies in vivo for inhibition of virus replication by modeling viral dynamics. Submitted for publication.
- Herz, A.V.M., Bonhoeffer, S., Anderson, R.M., May, R.M., and Nowak, M.A. (1996). Viral dynamics in vivo: Limitations on estimates of intracellular delay and virus decay. *Proc. Natl. Acad. Sci. USA* **93**, 7247-7251.
- Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., and Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123-126.
- Kirschner, D.E. and Perelson, A.S., A Model for the Immune System Response to

- HIV: AZT Treatment Studies, in *Mathematical Population Dynamics: Analysis of Heterogeneity, Vol.1, Theory of Epidemics*, O. Arino, D. Axelrod, M. Kimmel, and M. Langlais, Eds., 1995, pp. 295-310.
- Luzuriaga K, Mcmanus M, Britto P, Wu H, Lindsey J, Smith B, Mofenson L, Sullivan JL. Early combination antiretroviral therapy of vertical HIV-1 infection: Viral kinetics and control of viral replication. *The 5th Conference on Retroviruses and Opportunistic Infections*. Abstract #LB12.
- Melvin, A.J., Rodrigo, A.G., Hughes, J. Mohan, K., Manns-Arcuino, L., Coombs, R.W., Holte, S., Hagearty, P., Mullins, J.I., and Frenkel, L.M. (1998). Viral dynamics in children and infants undergoing highly active antiretroviral therapy (HAART). *The 5th Annual HIV Dynamics & Evolution Conference*, Sante Fe, New Mexico, USA.
- Merrill, S. (1987). AIDS: background and the dynamics of the decline of immunocompetence. In *Theoretical Immunology*, Part 2, A.S. Perelson, Ed., Addison-Wesley, Redwood City, Calif.
- Perelson, A.S. (1989). Modeling the interaction of the immune system with HIV. in *Mathematical and Statistical Approaches to AIDS Epidemiology* (Lect. Notes Biomath., Vol. 83), C. Castillo-Chavez, Ed., Springer-Verlag, New York.
- Perelson, A.S., Kirschner, D.E. and Boer R.D. (1993). Dynamics of HIV infection of CD4⁺ T cells. *Mathematical Biosciences* **114**, 81-125.
- 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.
- Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., and Ho, D.D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* **387**, 188-191.
- Phillips, A.N. (1996). Reduction of HIV concentration during acute infection: Inde-

- pendence from a specific immune response. *Science* **271**, 497-499.
- Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., Saag, M.S., and Shaw, G.M. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117-122.
- Wu, H. and Ding, A.A. (1998). Population HIV-1 Dynamics in Vivo: Applicable Models and Inferential Tools for Virological Data from AIDS Clinical Trials. *Biometrics*, in press.
- Wu, H., Ding, A.A., and DeGruttola, V. (1998a). Estimation of HIV Dynamic Parameters. *Statistics in Medicine*, in press.
- Wu, H., Kuritzkes, D.R., Clair, M.S., Spear, G., Connick, E., Landay, A., and Lederman, M.M. (1998b). Characterizing individual and population viral dynamics in HIV-1-infected patients with potent antiretroviral therapy: correlations with host-specific factors and virological endpoints. *The 12th World AIDS Conference*. Geneva, Switzerland.
- Wu, H., Ding, A.A., Sullivan, J.L., Luzuriaga, K. and ACTG 356 Investigators (1998c). The Effect of treatment efficacy and treatment time on HIV-1 dynamics. *The 5th Annual HIV Dynamics & Evolution Conference*, Sante Fe, New Mexico, USA.
- Wu, H., Kuritzkes, D.R., Clair, M.S., Kessler, H., Connick, E., Landay, A., Heath-Chiozzi, M., Rousseau, F., Fox, L., Spritzler, J., Leonard, J.M., McClernon, D.R., and Lederman, M.M. (1997). Interpatient variation of viral dynamics in HIV-1 infection: Modeling results of AIDS Clinical Trials Group Protocol 315. *The First International Workshop on HIV Drug Resistance, Treatment Strategies and Eradication*, Abstract 99, 66-67.
- Zhang, L., Dailey, P.J., He, T., et al. (1998). Rate of SIV particle clearance in Rhesus Macaques. *The 5th Conference on Retroviruses and Opportunistic Infections*. Abstract #549.

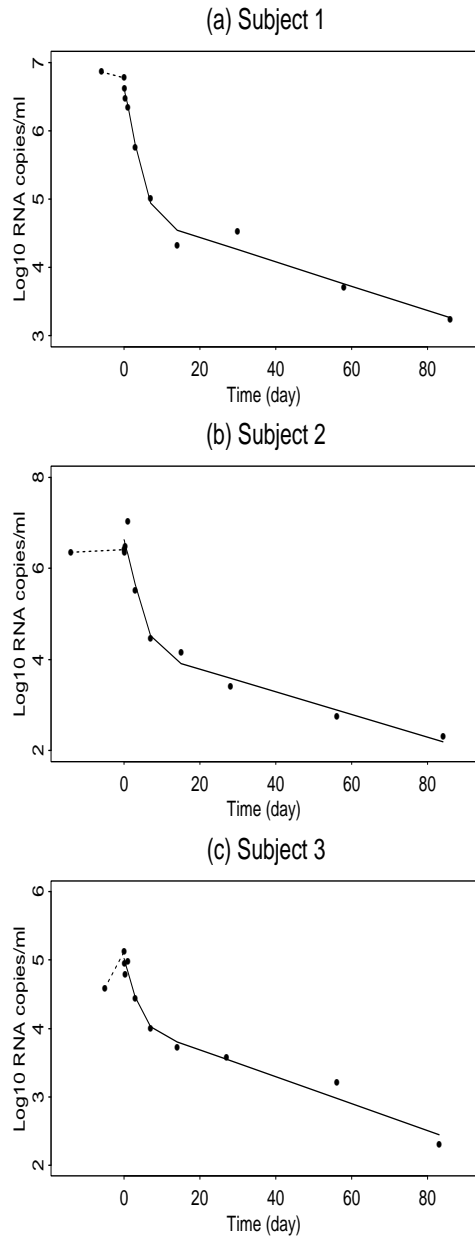


Figure 1: The plasma viral load data (\log_{10} scale) and fitted curves for three subjects from ACTG 356 using the suggested methods. Dots denote the observed data, solid lines denote the fitted curves, dots connected by dash lines are pre-treatment data.

Table 2: Simulation results of different estimation methods based on 500 runs. Initial Production/Clearance Ratio, $R = 0.5$. The numbers reported are percentages.

c	Schedule	Methods	Conv rate	d1					d2				
				Bias	STD	S.E.			Bias	STD	S.E.		
						n=1	n=16	n=50			n=1	n=16	n=50
2	Ideal	Naive	100	-14.3	16.4	21.7	14.9	14.5	-11.3	44.5	46	15.9	12.9
		WD	100	1.4	28	28	7.1	4.2	-5.4	45.1	45.5	12.5	8.4
		PNSS	96	6.1	32.3	32.9	10.1	7.6	-3.4	43.9	44.1	11.5	7.1
		PSS	99.6	-4.8	21.4	21.9	7.2	5.7	-6.9	44.3	44.7	13	9.3
	ACTG 356	Naive	98.8	-22.5	18.4	29	23	22.6	-2	13.2	13.4	3.9	2.7
		WD	98	-5.4	34.7	35.2	10.2	7.3	-1.8	13.6	13.8	3.8	2.6
		PNSS	88.8	-2.5	33.7	33.7	8.8	5.4	-1.2	14	14	3.7	2.3
		PSS	99.2	-11.9	25.1	27.8	13.5	12.4	-1.2	13.2	13.2	3.5	2.2
	ACTG 315	Naive	99.4	-25	22.1	33.3	25.6	25.2	-6	23.4	24	8.4	6.9
		WD	59.2	-25.4	31.2	40.2	26.6	25.8	-9.6	25.8	27.4	11.6	10.3
		PNSS	27.6	-5.3	25.2	25.6	8.2	6.4	-4	20.4	20.8	6.5	4.9
		PSS	98.6	-12	31.9	34	14.4	12.8	-4.8	23.4	23.8	7.6	5.8
3	Ideal	Naive	100	-10.7	18.1	21.2	11.6	11	-9.2	43.1	44.1	14.2	11
		WD	99.4	5.5	32.5	33	9.8	7.2	-5	43.9	44.1	12.1	8
		PNSS	94.8	9.6	35.2	36.5	13	10.8	-3.8	43.5	43.7	11.5	7.2
		PSS	99.8	1	25.2	25.2	6.4	3.7	-4.6	43.1	43.3	11.7	7.6
	ACTG 356	Naive	99.6	-15.2	21.7	26.5	16.1	15.5	-1.2	13	13	3.5	2.2
		WD	98.2	1.6	40.1	40.1	10.2	5.9	-1.4	13.8	13.8	3.7	2.4
		PNSS	87	5.1	36.1	36.4	10.4	7.2	-0.8	13.2	13.2	3.4	2
		PSS	99.6	-2.1	30.1	30.2	7.8	4.7	-0.6	12.8	12.8	3.3	1.9
	ACTG 315	Naive	98.4	-20.4	25.7	32.8	21.4	20.7	-5.2	23.2	23.8	7.8	6.1
		WD	56.4	-30.5	32.5	44.5	31.6	30.8	-10.2	27	28.8	12.2	10.9
		PNSS	22	-9.4	24.6	26.2	11.2	10	-3.6	20	20.2	6.2	4.6
		PSS	98.8	-4.6	37.4	37.6	10.4	7	-4.2	23.2	23.6	7.2	5.3
6	Ideal	Naive	99.8	-5.1	21	21.6	7.3	5.9	-7	42.4	42.8	12.7	9.2
		WD	99.6	9.7	38.4	39.6	13.6	11.1	-4.6	43.4	43.6	11.8	7.7
		PNSS	90	14.9	40.7	43.3	18	16	-2.8	42	42	10.9	6.6
		PSS	99.8	9.7	32	33.4	12.6	10.7	-2.6	42.2	42.2	10.9	6.5
	ACTG 356	Naive	99.4	-6.8	26	26.8	9.4	7.7	-0.8	12.8	12.8	3.3	2
		WD	97.4	4.9	44.8	45	12.2	8	-1.4	13.8	14	3.7	2.4
		PNSS	85.8	11.6	39.6	41.2	15.3	12.9	-1	13	13	3.4	2.1
		PSS	99.6	11.4	49	50.2	16.7	13.3	0	12.8	12.8	3.2	1.8
	ACTG 315	Naive	97.4	-12.6	30.2	32.8	14.7	13.3	-4.4	23	23.4	7.2	5.5
		WD	55.6	-35.6	28.5	45.6	36.3	35.8	-11.8	31.8	34	14.2	12.6
		PNSS	17.2	-10.3	22.5	24.6	11.7	10.8	-5.2	19.6	20.2	7.1	5.9
		PSS	97.8	8.3	44.4	45.1	13.9	10.4	-3.4	22.8	23	6.6	4.7

Table 3: Simulation results of different estimation methods based on 500 runs. Initial Production/Clearance Ratio, $R = 1$. The numbers reported are percentages.

c	Schedule	Methods	Conv rate	d1					d2				
				Bias	STD	S.E.			Bias	STD	S.E.		
						n=1	n=16	n=50			n=1	n=16	n=50
2	Ideal	Naive	99.6	-16.5	15.3	22.5	16.9	16.6	-13.3	45.5	47.3	17.5	14.8
		WD	100	0	26	26	6.5	3.7	-5.7	46.1	46.3	12.9	8.7
		PNSS	97.2	4.1	29.8	30.1	8.5	5.9	-4	45.7	45.9	12.1	7.6
		PSS	99	-8	19.5	21.1	9.4	8.5	-8.1	45.1	45.7	13.9	10.3
	ACTG 356	Naive	98.8	-26	16.7	30.9	26.3	26.1	-2.4	13.4	13.6	4.1	3.1
		WD	98.6	-7.1	34.2	34.9	11.1	8.6	-2	14	14	4	2.8
		PNSS	89.8	-6.2	29	29.7	9.5	7.4	-1.2	13.4	13.4	3.6	2.2
		PSS	99.4	-17.1	22.4	28.2	18	17.4	-1.8	13.4	13.6	3.8	2.6
	ACTG 315	Naive	99	-27.2	20.2	33.8	27.7	27.3	-6.6	23.6	24.4	8.9	7.4
		WD	60.6	-23.1	28.6	36.6	24.2	23.5	-9.2	25.4	27	11.2	9.9
		PNSS	30.4	-4.7	22.2	22.6	7.3	5.7	-3.2	21.2	21.4	6.2	4.4
		PSS	98.4	-16	28.8	32.9	17.5	16.5	-5.4	23.2	23.8	7.9	6.3
3	Ideal	Naive	100	-13.6	16.9	21.7	14.2	13.8	-10.8	43.7	45.1	15.4	12.4
		WD	99.8	4	30	30.2	8.5	5.8	-4.8	44.5	44.7	12.1	7.9
		PNSS	96.8	8.2	33.2	34.1	11.7	9.4	-3.6	44.3	44.3	11.6	7.2
		PSS	99.8	-3.3	22.7	22.8	6.6	4.6	-6.2	43.7	44.1	12.6	8.8
	ACTG 356	Naive	99.4	-19.6	19.1	27.4	20.2	19.8	-1.8	13	13.2	3.7	2.6
		WD	99	-0.3	37.4	37.4	9.4	5.3	-1.4	13.8	13.8	3.7	2.4
		PNSS	88.8	2.2	33.9	34	8.8	5.3	-1	13.2	13.2	3.4	2.1
		PSS	98.8	-7.9	26.5	27.6	10.3	8.7	-1	13	13	3.4	2.1
	ACTG 315	Naive	99.6	-23.9	23.2	33.3	24.6	24.1	-5.8	23.4	24	8.2	6.7
		WD	59	-26.4	31.7	41.3	27.6	26.8	-9.8	26.2	28	11.8	10.5
		PNSS	27.2	-5.4	28.5	28.8	8.9	6.7	-4.6	22	22.4	7.2	5.6
		PSS	98.4	-10.3	33.8	35.3	13.3	11.4	-4.8	23.2	23.6	7.5	5.8
6	Ideal	Naive	99.8	-8.4	19.4	21.1	9.7	8.8	-8.4	42.8	43.6	13.6	10.4
		WD	99.4	8.4	35.5	36.5	12.2	9.8	-4.6	43.6	43.8	11.8	7.7
		PNSS	94.6	12.6	37.6	39.6	15.7	13.7	-3.4	43.4	43.6	11.4	7
		PSS	99.6	4.8	28.5	28.9	8.6	6.3	-4	42.8	42.8	11.4	7.3
	ACTG 356	Naive	100	-11.3	23.7	26.2	12.8	11.8	-1	13	13	3.4	2.1
		WD	98.4	5.8	44.9	45.2	12.6	8.6	-1.2	13.8	13.8	3.7	2.3
		PNSS	88	9.7	36.9	38.1	13.4	11	-0.6	13.2	13.2	3.4	2
		PSS	99.8	4.2	39.6	39.8	10.8	7	-0.4	12.8	12.8	3.2	1.9
	ACTG 315	Naive	99.6	-17.1	27.7	32.5	18.4	17.5	-4.6	23.2	23.6	7.4	5.7
		WD	56.2	-32.7	29.1	43.7	33.5	33	-10.8	29	30.8	13	11.6
		PNSS	21	-10.3	24.6	26.6	12	10.9	-3.6	20.4	20.6	6.2	4.6
		PSS	97.8	0.6	41	41	10.3	5.8	-3.8	23.2	23.4	6.9	5

Table 4: Simulation results of different estimation methods based on 500 runs. Initial Production/Clearance Ratio, $R = 3$. The numbers reported are percentages.

c	Schedule	Methods	Conv rate	d1					d2				
				Bias	STD	S.E.			Bias	STD	S.E.		
						n=1	n=16	n=50			n=1	n=16	n=50
2	Ideal	Naive	99.6	-22.3	12.9	25.7	22.5	22.4	-20.8	48.8	53.1	24.1	21.9
		WD	100	-3.2	20.9	21.2	6.1	4.4	-6.9	48.6	49	14	9.7
		PNSS	94.8	0.7	23.5	23.5	5.9	3.4	-3.5	47.6	47.8	12.4	7.6
		PSS	90.8	-14.7	14.6	20.8	15.1	14.8	-12.7	47.8	49.4	17.4	14.4
	ACTG 356	Naive	99.4	-34.2	13.4	36.8	34.4	34.3	-3.2	22.3	22.5	6.4	4.5
		WD	98.2	-10.5	28.9	30.7	12.7	11.3	-2.4	15.3	15.3	4.5	3.2
		PNSS	90.8	-14.1	22.8	26.9	15.2	14.5	-2	14	14.2	4	2.8
		PSS	96	-27.9	16.5	32.4	28.2	28	-3.2	14	14.2	4.7	3.8
	ACTG 315	Naive	99	-31.3	15.9	35.2	31.6	31.4	-8.2	24.1	25.5	10.2	8.9
		WD	67.6	-12.4	28.6	31.2	14.3	13	-7.6	24.7	25.9	9.8	8.4
		PNSS	37.6	3.5	23.9	24.1	6.9	4.9	-2.8	21.5	21.7	6.1	4.1
		PSS	96.4	-23.5	21.4	31.8	24.1	23.7	-7	23.9	24.9	9.2	7.8
3	Ideal	Naive	99.6	-20.4	14	24.8	20.7	20.5	-17.1	46.3	49.3	20.6	18.3
		WD	99.6	1	24.3	24.3	6.2	3.6	-5	46.5	46.7	12.7	8.3
		PNSS	98.4	4.6	28.2	28.6	8.4	6.1	-3.4	46.1	46.1	12	7.4
		PSS	97.2	-12.5	17.1	21.2	13.2	12.7	-10.7	45.7	46.9	15.7	12.5
	ACTG 356	Naive	99.2	-29.9	15	33.5	30.1	30	-3.4	13.6	14	4.8	3.9
		WD	98.8	-3.2	32.3	32.4	8.7	5.6	-1.4	13.6	13.6	3.7	2.4
		PNSS	94	-6.1	26.9	27.6	9.1	7.2	-1.2	13.4	13.4	3.6	2.2
		PSS	98	-21.9	19.3	29.2	22.4	22.1	-2.4	13.6	13.8	4.2	3.1
	ACTG 315	Naive	98	-31	17.7	35.7	31.3	31.1	-7.8	23.9	25.1	9.8	8.5
		WD	64.2	-16.9	30.7	35	18.6	17.4	-8.6	24.7	26.1	10.6	9.3
		PNSS	36	-1.2	26.5	26.4	6.7	3.9	-4.6	21.5	22.1	7.1	5.5
		PSS	97	-22.2	24.6	33.2	23	22.5	-6.6	23.7	24.5	8.9	7.4
6	Ideal	Naive	99.8	-16.4	16	22.9	16.9	16.6	-13.8	44.7	47	17.8	15.2
		WD	100	6.3	29.2	29.8	9.6	7.5	-3.4	45.3	45.5	11.8	7.3
		PNSS	98.6	10.4	33.7	35.3	13.4	11.4	-2.4	45.5	45.5	11.6	6.9
		PSS	99	-7.1	21	22.1	8.8	7.7	-8.2	44.5	45.1	13.8	10.3
	ACTG 356	Naive	98.6	-23.2	17.3	28.9	23.6	23.3	-2.6	13.4	13.6	4.2	3.2
		WD	99.2	4.6	38.1	38.4	10.6	7.1	-1	13.4	13.4	3.5	2.1
		PNSS	94.4	4.1	30.9	31.1	8.7	6	-0.8	13.2	13.2	3.4	2
		PSS	99.4	-12.3	24.1	27.1	13.7	12.8	-1.6	13.2	13.4	3.7	2.5
	ACTG 315	Naive	99.4	-27.6	21.3	35	28.1	27.8	-6.8	23.7	24.7	9	7.6
		WD	61.4	-22.8	28.8	36.6	23.9	23.2	-9.2	25.7	27.3	11.2	9.9
		PNSS	31.8	-3.9	23.5	23.7	7.1	5.1	-4.8	21.5	22.1	7.2	5.7
		PSS	98.2	-16.5	30.6	34.6	18.2	17.1	-5.4	23.3	23.9	7.9	6.3

Table 5: The estimates of d_1 and d_2 for three subjects using the four methods.

Sub- ject	R	\hat{d}_1				\hat{d}_2			
		NAIVE	WD	PNSS	PSS	NAIVE	WD	PNSS	PSS
1	$R < 1$	0.673	0.624	0.594	0.709	0.041	0.041	0.040	0.041
2	$R = 1$	0.763	1.750	1.043	0.830	0.058	0.064	0.060	0.058
3	$R > 1$	0.549	0.794	0.563	0.598	0.045	0.046	0.045	0.045