Published in final edited form as:

Stat Methods Med Res. 2007 August ; 16(4): 365–373. doi:10.1177/0962280206071850.

Segmental Modeling of Changing Viral Load to Assess Drug Resistance in HIV Infection¹

H Liang

Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, USA

Abstract

Studies of viral dynamics are important to our understanding of the pathogenesis of HIV-1 infection and in assessment of the potency of antiretroviral therapies. Although many different viral dynamic models and methods for estimating viral dynamic parameters have been proposed and used in various studies, none has been entirely satisfactory. We propose here a segmental model to describe the viral load, and estimate the dynamic parameters by using a mixed-effects model. We address the relation between the baseline viral load and the decay rate of the first phase of viral load decay, and divide patients into three categories on the basis of their changing viral load patterns.

1 Introduction

In clinical studies of the efficacy of anti-HIV drugs, the changing viral load in subjects' blood can be approximated by an exponential function within the first 2 weeks after starting the treatment. In other words, the log of the viral load decreases linearly and sharply with time (i.e., has a large negative slope)^{1, 2}. After the first 2 weeks, however, the picture is very different and shows multiple phases of decline and increase. Within the first 2-3 months but after the first 2 weeks, the relation between the log of the viral load and time remains linear, but the curve is flat (i.e., horizontal) while the virus is released by macrophages and other long-lived infected cells³. At some time between the third and eighth month, the viral load in some individuals increases sharply and reaches the level measured before treatment, presumably because of the development of resistance of the virus to the drugs.

In studying viral dynamics to understand the pathogenesis of HIV-1 infection and to assess the potency of antiretroviral therapies, viral load decay rates have been estimated by using linear models in a log₁₀ scale in short-term (about 1 week) dynamics^{1, 2, 4-7} and biphasic models (nonlinear biexponential model) in longer term (2 to 8 weeks) dynamics^{3, 8-11}. Wu et al.¹² compared linear and biphasic nonlinear modeling and found that linear modeling may result in misleading conclusions because one has to truncate the data, but it is not clear where the data should be cut. The biphasic nonlinear modeling¹² showed promise. However, implementation of the biphasic nonlinear modeling becomes a challenge because the computation is very sensitive to initial values and convergence problems often occurs. We propose here a piecewise-linear segmental model as a compromise. This model does not require truncation of data and is more easily implemented than the biphasic nonlinear model.

¹Address for correspondence: Hua Liang, Department of Biostatistics and Computational Biology, University of Rochester Medical Center, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, USA..

2 Model and Methods

ACTG 315¹³ was developed by the Immunology Research Agenda Committee of the U.S. NIAID-sponsored ACTG. The protocol was conducted at several university hospitals such as Case Western Reserve University Lukes Medical Center and the University of Colorado Health Sciences Center.

Fig. 1 presents data from the ACTG 315 study, as a scatter plot of viral load (in $\log_{10}$ scale) on treatment days 0, 2, 7, 10, 14, 28, 56, 84, 168, and 336. Fig. 1 shows that the viral load in most patients declines rapidly within the first 2 weeks and the curve has a large negative slope; after 2 weeks, however, the trend of decline decreases and the curve becomes flat. At about 10 weeks, the viral loads rebound, and most increase until the end of treatment. To better understand HIV pathogenesis and to improve treatment management and care of patients with AIDS, it is important to identify when a patient's viral load declines and when the change to a slower decline and rebound occur.

2.1 Model

Fig. 1 shows that we may describe the effect of treatment on viral load in 3 patterns (Fig. 2): rapid decline followed by slow decline and rebound; rapid decline followed by slow decline; and decline followed by rebound. Rapid decline followed by slow decline (Fig 2b) indicates that drug resistance does not develop, and the treatment can be thought successful. We are interested in the rates of decline and rebound and the time points at which changes occur. From the biological perspective, one may link the first change point to the time at which the macrophages or other long-lived HIV-infected cells from lymphoid tissues release free HIV into the blood [3]; one may link the second change point to the time at which the virus has developed drug resistance. These time points have important clinical implications: they signal the time to change drugs to avoid the problem of drug resistance.

We use a unified function to describe the patterns in Fig. 2. The form for (a) is derived first because (b) and (c) are its special cases. Let t_L and t_R denote the times of the first and second change points respectively. The first segment of the curve is described by $\eta_0 + \eta_1 t$, represents a rapid decline in viral load after treatment. η_0 represents baseline viral load and η_1 represents the rate of viral load change. The second line is described by $\eta_2 + \eta_3 t$, and represents a less rapid decline after some unknown time point t_L . This slower decline occurs as a result of the release of HIV from macrophages or other long-lived cells from the lymphoid tissues. The third segment is described by $\eta_4 + \eta_5 t$ and represents an increase in the viral load after some unknown time point t_R as a result of drug resistance. Notice that the first two segments agree at t_L , while the last two segments agree at t_R . Denote $\Delta_1 = \eta_3 - \eta_1$ and $\Delta_2 = \eta_5 - \eta_3$. By combining the three segments we can describe the pattern of (a) in Fig 2 by the form

$$\eta_0 + \eta_1 t + \Delta_1 (t - t_L)_+ + \Delta_2 (t - t_R)_+ \quad (1)$$

where s_+ denotes $\max(s, 0)$ for a variable s . When $\eta_1 < 0$, $\eta_1 + \Delta_1 < 0$, and $\eta_1 + \Delta_1 + \Delta_2 > 0$, (1) represents (a); when $\eta_1 < 0$, $\eta_1 + \Delta_1 < 0$, and $\Delta_2 = 0$, (1) represents (b); and when $\eta_1 < 0$, $\eta_1 + \Delta_2 > 0$, and $\Delta_1 = 0$, (1) represents (c).

Suppose that there are n HIV-infected individuals. Let y_{ij} be the viral load $\log_{10}(\text{RNA})$ of the i th patient at time t_{ij} . Because of between-subject and within-subject variations, the parameters η_0 , η_1 , Δ_1 , Δ_2 , t_L , and t_R vary with respect to different patients. We model $\log_{10}(\text{RNA})$ by the mixed-effects model¹⁵:

$$y_{ij} = \eta_{0i} + \eta_{1i}t_{ij} + \Delta_{1i}(t_{ij} - t_{Li})_+ + \Delta_{2i}(t_{ij} - t_{Ri})_+ + \varepsilon_{ij} \quad (2)$$

where ε_{ij} is the random measurement error in measuring y_{ij} . It is assumed that the ε_{ij} 's are independently distributed normal random variables with means 0 and variance σ_i^2 independently of $\{\eta_{0i}, \eta_{1i}, \Delta_{1i}, \Delta_{2i}, t_{Li}, t_{Ri}\}$ for all $i = 1, \dots, n$. That is, $\varepsilon_{ij} \sim N(0, \sigma_i^2)$. $\eta_{0i}, \eta_{1i}, \Delta_{1i}, \Delta_{2i}, t_{Li}$, and t_{Ri} are also assumed to be independently normally distributed with mean parameters $\eta_0, \eta_1, \Delta_1, \Delta_2, t_L$, and t_R ; i.e., $\eta_{0i} = \eta_0 + b_{0i}$ for $i = 1, \dots, n$ and similarly for $\eta_{1i}, \Delta_{1i}, \Delta_{2i}, t_{Li}$, and t_{Ri} . b_{0i} indicates the individual variation. Our interest is to estimate both population parameters $\eta_0, \eta_1, \Delta_1, \Delta_2, t_L, t_R$ and individual parameters $\eta_{0i}, \eta_{1i}, \Delta_{1i}, \Delta_{2i}, t_{Li}, t_{Ri}$. These parameters can be estimated by existing algorithms such as that of Lindstrom and Bates¹⁶, and implemented can be gained by using the Splus/R function nlme¹⁵.

2.2 Relation between baseline RNA level and first decay rate

The first decay rate η_1 indicates the potency of antiviral therapies^{10, 14}. By knowing the potency of a treatment for a particular patient at an early stage, we may be able to avoid giving that patient the less potent regimens. In addition, if we know that there is a correlation between the baseline RNA viral load and the rate of the first phase of decay (the indicator of treatment potency), we may be able to use the baseline viral load to determine whether a treatment regimen is potent enough for a particular patient. This correlation will help clinicians to select a treatment for their patients. Drugs may show different potencies for different patients being treated on the same regimen, because patients may absorb the drug differently, and patients' immune systems and other factors related to their responses to the drugs may be different.

3 Data Analysis

We fit the viral load trajectory (viral dynamics) data by the model (2), with response variables being the numbers of HIV-1 RNA copies per mL of plasma on a $\log_{10}$ scale. We take 14 and 70 as initial values of t_L and t_R . Using the data within one of time intervals $[0, 14]$, $(14, 70]$, and $(70, 336]$ to fit a simple linear model, and the estimated coefficients are taken as initial values for modeling (2). The population estimates of the parameters of model (2) are $\eta_0 = 5.066$ (s.e. 0.2623), $\eta_1 = -0.127$ (s.e. 0.033), $\Delta_1 = 0.139$ (s.e. 0.035), $\Delta_2 = 0.019$ (s.e. 0.013), $t_L = 15$ (s.e. 4.28), and $t_R = 78$ (s.e. 19.08). In a consequence, our population 3-segment curve is $5.066 - 0.127t$ when $t \leq 15$; $3.341 - 0.012t$ when $15 < t \leq 78$; and $2.093 + 0.004t$ when $t > 78$. On average the period from entry time to 15 days can be taken as the first stage and the period from 15 days to 78 days can be taken as the second stage. The individual parameters $\eta_{0i}, \eta_{1i}, \Delta_{1i}, \Delta_{2i}, t_{Li}$, and t_{Ri} are obtained by plugging in empirical Bayesian estimates for the random effects, and the individual curves are obtained consequently. Estimates of the variances of the random effects associated to $\eta_0, \eta_1, \Delta_1, \Delta_2, t_L, t_R$ are 0.02, 0.05, 0.05, 0.004, 4.67, 8.29.

We found that two patients' data were too irregular to be described by any of the three kinds of curves, but the patterns of changing viral load of 8, 26, and 12 patients followed the patterns shown in Fig 2 (a), (b), and (c), respectively. The changing viral load patterns of 9 patients are shown in Fig 3. The population and individual curves are indicated by dotted and solid lines, respectively, and the observed data are indicated by plus signs. The first row of curves shows three-segment patterns like Fig. 2(a). The second row of curves shows two-segment patterns like Fig. 2(b). The third row of curves shows two-segment patterns like Fig. 2(c). For each patient, the estimated individual curve is much closer than the population curve to the observed values. This finding reflects the benefit of mixed-effects models in that the between-subject and within-subject variations are simultaneously accounted for. The population curve indicates that the viral load declines rapidly at the beginning of treatment, maintains a shallow decline

from day 15 to day 78, then slightly rebounds. All individual curves are similar to the population curve at the beginning of treatment, though there are differences in magnitude. However, after 15 days the patterns of the individual curves are diversely different from the population curve. Even within a row, the slopes or change points for different patients have different magnitudes. The between-row difference is not only in magnitudes, but also in patterns. The virus did not develop resistance to drug in more than half the patients.

A scatter plot (Fig. 4) of the baseline viral load and the first phase decay rates, $\hat{\eta}_{1i}$, which are estimated using model (2), shows that the relation between these two quantities is surprisingly unlike the traditional linear relation, being positive when $\log_{10}(\text{RNA}) < 4.5$ and negative when $\log_{10}(\text{RNA}) > 4.5$. The line in Fig. 4 presents an intuitive impression of this relation.

To further explore this finding, we separated patients into two groups on the basis of their baseline viral load being less than 4.5 or not. The changing viral load patterns of those groups between initial treatment and day 14 are shown in Fig. 5. Some patients' viral loads dramatically declined (left plot) and showed more promise than did those of patients whose baseline viral load was greater than 4.5 (right plot).

4 Discussion

A major obstacle to the treatment of HIV-infection with antiretroviral drugs is the development of drug resistance. To avoid this problem, we need to be able to estimate the time at which resistance will develop. To perform this estimation, we have developed an approach that assumes a segmental linear model with unknown change points for the viral load as indicated by $\log_{10}(\text{RNA})$. In this model, the first change point is the time at which the macrophages or other long-lived cells release HIV into the blood from lymphoid tissues³ or the time at which patients develop drug resistance; the second change point is the time at which the virus develops resistance to the drugs.

For each patient in the ACTG 315 study, we knew the patient's viral load (copies of HIV-1 RNA per *ml* of blood) at various time points after treatment. Using these data and log transformation, we estimated that the first change point would overall occur 15 days after the start of treatment. The population estimates of the slopes of the three-segmental lines were -0.12 , -0.012 , and 0.004 respectively. The standard errors of these estimates are given by 0.046 , 0.01 , and 0.003 . These estimates indicate that most patients had the similar slopes for the first segment but not the other segments. This observation may be explained by noting that for some patients, the second segment was very short or barely noticeable, while for other patients, the second segment was long, and that although drug resistance developed in some patients, other patients did not develop or had barely noticeable drug resistance. The second and third segments of the changing viral load curves therefore showed considerable variation between individuals. Some patients' changing viral load patterns may not follow the two or three-segment pattern; a more complex form may be required to describe these patterns. Challenges of implementation and biological interpretation then arise.

Model (2) may be extended to a more general form, which contains covariates such as CD4+ cell counts. The study of the model is interesting and requires additional efforts, but it is beyond the scope of this paper.

Acknowledgments

The author is grateful to one referee for helpful suggestions. This research was partially supported by NIH/NIAID grants AI62247-01, AI59773, and AI50020.

References

1. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 1995;373:123–6. [PubMed: 7816094]
2. Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, Lifson JD, Bonhoeffer S, Nowak MA, Hahn BH, Saag MS, Shaw GM. Viral dynamics in human immunodeficiency virus type-1 infection. *Nature* 1995;373:117–22. [PubMed: 7529365]
3. Perelson AS, Essunger P, Cao Y, Vesanen M, Hurley A, Saksela K, Markowitz M, Ho DD. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* 1997;387:188–91. [PubMed: 9144290]
4. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 1996;271:1582–6. [PubMed: 8599114]
5. Polis MA, Sidorov IA, Yoder C, et al. Correlation between reduction in plasma HIV-1 RNA concentration 1 week after start of antiretroviral treatment and longer-term efficacy. *The Lancet* 2001;358:1760–5.
6. Louie, M.; Hogan, C.; Hurley, A., et al. Determining relative efficacy of tenofovir DF using frequent measurements of HIV-1 RNA during a short course of monotherapy in antiretroviral drug naive individuals; 9th Conference on Retroviruses and Opportunistic Infections; Seattle, WA. February 24–28, 2002; Abstract# 3
7. Sankatsing, S.; Weverling, G.; van't Klooster, G., et al. TMC125 Monotherapy for 1 week results in a similar initial rate of decline of HIV-1 RNA as therapy with a 5-drug regimen; 9th Conference on Retroviruses and Opportunistic Infections; Seattle, WA. February 24–28, 2002; Abstract# 5
8. Notermans DW, Goudsmit J, Danner SA, de Wolf F, Perelson AS, Mittler J. Rate of HIV-1 decline following antiretroviral therapy is related to viral load at baseline and drug regimen. *AIDS* 1998;12:1483–90. [PubMed: 9727569]
9. Wu H, Kuritzkes DR, McClernon DR, et al. Characterization of viral dynamics in human immunodeficiency virus type 1-infected patients treated with combination antiretroviral therapy: relationships to host factors, cellular restoration, and virologic end points. *Journal of Infectious Diseases* 1999;179:799–807. [PubMed: 10068574]
10. Ding AA, Wu H. Relationships between antiviral treatment effects and biphasic viral decay rates in modeling HIV dynamics. *Mathematical Biosciences* 1999;160:63–82. [PubMed: 10465932]
11. Wu H, Ding AA. Population HIV-1 dynamics in vivo: applicable models and inferential tools for virological data from AIDS clinical trials. *Biometrics* 1999;55:410–8. [PubMed: 11318194]
12. Wu H, Zhao CX, Liang H. Comparison of linear, nonlinear and semiparametric mixed-effects models for estimating HIV dynamic parameters. *Biometrical Journal* 2004;46:233–45.
13. Lederman MM, Connick E, Landay A, et al. Immunologic responses associated with 12 weeks of combination antiretroviral therapy consisting of zidovudine, lamivudine, and zalcitabine: results of AIDS Clinical Trials Group Protocol 315. *Journal of Infectious Diseases* 1998;178:70–9. [PubMed: 9652425]
14. Perelson AS, Nelson PW. Mathematical analysis of HIV-1 dynamics in vivo. *SIAM Review* 1999;41:3–44.
15. Pinheiro, JC.; Bates, DM. *Mixed-Effects Models in S and S-PLUS*. Springer; New York: 2000.
16. Lindstrom MJ, Bates DM. Nonlinear mixed-effects models for repeated measures data. *Biometrics* 1990;46:673–87. [PubMed: 2242409]

Figure 1.
Viral load data for patients in the ACTG 315 study.

Figure 2.

Illustrative plots of 3 typical viral load trajectories: (a) 3-segment curve consisting of rapid and slow decline followed rebound; (b) 2-segment curve consisting of rapid and slow decline; and (c) 2-segment curve consisting of decline and rebound.

Figure 3.

Individual estimates of viral load obtained by using model (2). The observed values are indicated by plus signs. The population and individual curves are indicated by broken and solid lines, respectively. The first row of curves follows the pattern shown in Fig. 2(a), the second row follows Fig. 2(b), and the last row follows Fig. 2(c).

Figure 4. Scatterplot of the estimated viral load decay rates $\hat{\eta}_{1i}$ against baseline viral load. The curve was estimated by smoothing spline.

Figure 5.

Viral load trajectories of two groups of patients: on the left, those whose baseline viral load are less than 4.5, and on the right, patients whose baseline viral load was higher than 4.5.