Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mathematical Biosciences 200 (2006) 1–27
Mathematical Biosciences
www.elsevier.com/locate/mbs

Estimating kinetic parameters from HIV primary infection data through the eyes of three different mathematical models

M.S. Ciupe [a], B.L. Bivort [b], D.M. Bortz [c], P.W. Nelson [a,c,*]

[a] *Los Alamos National Lab, MS K710, Los Alamos, NM 87545, United States*
[b] *Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, United States*
[c] *Math Biology Research Group, Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, MI 48109, United States*

Received 24 September 2003; received in revised form 12 December 2005; accepted 18 December 2005
Available online 9 February 2006

Abstract

The dynamics of HIV-1 infection consist of three distinct phases starting with primary infection, then latency and finally AIDS or drug therapy. In this paper we model the dynamics of primary infection and the beginning of latency. We show that allowing for time delays in the model better predicts viral load data when compared to models with no time delays. We also find that our model of primary infection predicts the turnover rates for productively infected T cells and viral totals to be much longer than compared to data from patients receiving anti-viral drug therapy. Hence the dynamics of the infection can change dramatically from one stage to the next. However, we also show that with the data available the results are highly sensitive to the chosen model. We compare the results using analysis and Monte Carlo techniques for three different models and show how each predicts rather dramatic differences between the fitted parameters. We show, using a χ^2 test, that these differences between models are statistically significant and using a jackknifing method, we find the confidence intervals for the parameters. These differences in parameter estimations lead to widely varying conclusions about HIV pathogenesis. For instance, we find in our model with time delays the existence of a Hopf bifurcation that leads to sustained oscillations and that these

* Corresponding author. Address: Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, MI 48109, United States. Tel.: +1 (734) 763 3408.
E-mail address: pwn@umich.edu (P.W. Nelson).

0025-5564/$ - see front matter
doi:10.1016/j.mbs.2005.12.006

oscillations could simulate the rapid turnover between viral strains and the appropriate CTL response necessary to control the virus, similar to that of a predator–prey type system.

Keywords: HIV; Primary infection; Time delays; Monte Carlo; Jackknifing; Model sensitivity

1. Introduction

The dynamics of HIV-1 consist of three distinct phases starting with primary infection, then latency and finally AIDS or drug therapy. Primary infection begins when the host first becomes infected with HIV-1. During this period there is a rapid increase in the number of viral particles in the plasma which can reach well over 10 million copies per ml. While in primary infection the infectious HIV-1 particles seek out target cells, mostly $CD4^+$ T cells, and infect them until a peak in viral concentration is reached. After which, due to limitations of target cells and/or the emergence of cytotoxic (or effector) T cells that target HIV-1 infected cells, the viral load begins to decline. These two factors play a major role in the control of viral loads during primary infection.

Primary infection ends when the viral loads reach a set point which defines the beginning of latency. Viral load at set point is a good indicator of the future of the disease. During latency, there can be numerous changes in the viral load totals but not to the extent seen in primary infection. In many patients the behavior is characteristic of a damped oscillating system in which there are substantial oscillations in viral loads until a quasi-steady state level is reached. This state can last for many years until the virus becomes more active, i.e., the patient develops AIDS, or the patient receives an anti-viral therapy which causes viral totals to greatly diminish.

We examine the dynamics of primary infection and latency in this paper by considering a delay differential equation model, where the time delay represents the lag in effector cell activation after initial infection. Adding the dynamics of the effector cells to previously existing models increases the number of model parameters and so we employ a Monte Carlo numerical technique and jackknifing to estimate the parameters. We assume certain parameters to have the same value as given in Stafford et al. [1] while allowing for others, such as the growth and death rates of effector cells, the length of the time delay, and the death rates of productively infected T cells and virus to be estimated. Our analyses give widely varying estimates of these kinetic rates when compared to the estimates in Stafford et al. [1], such as the death rates of productively infected T cells and virus, while providing very similar dynamical fits to the patient data. These results show how sensitive the fits to patient data are to the chosen model, especially when the data is insufficient, i.e., not enough data points for the studied period to distinguish between the models. Even though the parameter estimates are sensitive to the model, each model allows for a different interpretation of the biology, and we show through the analysis of each model its usefulness in future research of HIV primary infection. For instance, we find for certain parameter ranges that the dynamics lead to sustained oscillations in viral loads which implies the viral dynamics are actively changing during latency. These dynamics are not seen in the models without the time delays unless certain non-linearities are introduced [2].

2. Basic model

The models in this paper seek to describe the dynamics of HIV-1 viral load during primary infection. Stafford et al. [1] proposed an ODE model that illustrates two possible immune responses that may control the steady state level of virus after primary infection. Their model consisted of three types of cells: target cells ($CD4^+$ T cells), T, productively infected $CD4^+$ T cells, T^*, and free virus, V. They assume target cells decay exponentially with no proliferation. Their model accounts for primary infection and latency differently then previous models [3–7] by including an exponential term to account for the cytotoxic activity of effector cells targeting infected cells. They showed that without this term their model was unable to account for the long term steady state viral loads seen in some of the patients. We pursued the idea of $CD8^+$ T cell reduction of productively infected T cells by developing an equation for CTL's, E.

2.1. Effector cell equation

Determining the proper equation for effector cells is not straightforward due to the complexities of the biology. Previous work by DeBoer and Perelson [2] and Nowak and Bangham [8] considered effector cell dynamics with

$$\frac{dE}{dt} = \alpha_E EI - d_E E, \tag{1}$$

where α_E is the rate with which an effector cell, E, comes in contact with a productively infected T cell, I, and d_E is the death rate of effector cells. Note that their notation assumes I instead of T^* for these cells. However, a problem with this choice for effector cells is that it leads to a steady state condition $I_{ss} = \frac{d_E}{\alpha_E}$ and thus the steady state dynamics of the productively infected T cells are independent of any viral parameters. Therefore, one needs to assume a steady state level of productively infected T cells and set the parameters for effector cells based on this value. This then puts a constraint on the full model's dynamics and thus limits its predictive ability.

One could also use a more sophisticated equation for effector cells [2] but this would only introduce more variability into the model and thus take away from the focus of simply including the dynamics for effector cells. We found that if we considered (1) in our model that the Monte Carlo simulations did not always settle to a global minima, especially when we fixed $c = 3\ \text{day}^{-1}$. Also, when we fixed the steady state levels as done in [2,8] we find using (1) that the models' predicted maximal levels of effector cells during primary infection were on average less then $0.124\ \mu\text{l}^{-1}$ and thus well below the values from the literature. Data from the literature shows the average number of effector cells to range between 375 and $1100\ \mu\text{l}^{-1}$ [9,10].

We chose to assume that the effector cells were activated at a rate proportional to the level of productively infected T cells. In our model, these cells become activated at a rate p and decay with a half life of $1/d_E$. We will also consider a time delay in the activation term to account for the time needed by effector cells to recognize an infection. For our case we chose,

$$\frac{dE}{dt} = pT^* - d_E E. \tag{2}$$

We consider the CTL burst size to be dependent on the duration of antigen presentation. Experimental studies have shown that following a brief antigen encounter $CD8^+$ T cells divide into

effector cells and the differentiation continues even after the antigenic stimulus has been removed. The capacity of CTLs to undergo prolonged division in the absence of antigen has been monitored in *Listeria monocytogenes* bacterial infection in vitro by Mercado et al. [11] and Wong and Pamer [12]. van Stipdonk et al. [13] emphasized this idea by engineering antigen presenting cells which regulate $CD8^+$ T cells priming. The results in [13] were used by Antia et al. [14] to argue that the simple predator-prey type dynamics for CTL cell responses was not appropriate and therefore the level of effector cells is not simply proportional to the concentration of productively infected T cells. However, van Stipdonk et al. [15] followed up their paper in 2001 and showed that $CD8^+$ T cells that received brief antigenic stimulation undergo apoptosis much faster than those that receive sustained signals from antigen presenting cells. Hence, disproving their earlier claims and again leaving the question, of what is the best equation for effector cells, unanswered.

Another issue to consider is the importance of costimulatory cytokines in the CTL proliferation pathway. Wong and Pamer et al. [12] used IL-2 to stimulate CTL differentiation after antigen removal, and although CTLs are capable of producing their own IL-2 [16], van Stipdonk has shown that a prolonged antigenic presence increases the capacity of CTLs to produce and respond to IL-2 simulations [15]. Finally, there are no studies that explain how the duration of antigenic stimulation regulates CTL responses in viral infections like HIV. One difference between the immune response in HIV infection and *Listeria* infection regards the number of $CD8^+$ T cells involved in the response. *Listeria* induces huge proliferation rates, while only 1.6–18.4% of the total number of $CD8^+$ T cells respond to HIV infections [17]. Also, the physiological relevance of this programmed immune response is limited by the fact that conclusions were determined by in vitro experiments. Hence, we are far from understanding the dynamics of CTL division cycles and further studies are needed to assess whether or not the immune response depends on the duration of antigenic exposure.

2.2. Effector model

We will consider both a non-delay model (effector model) and one with a time delay, $t-\tau$, in the activation of the effector cells (Delay model 1). In either case the steady state level of productively infected T cells is now controlled by viral parameters. We find in our case that when we consider (2) in the model that the effector cells can range from 1 to 980 μl^{-1} which is in the same range as found in [8] and also overlaps with the values given in [9,10]. With this in mind we have the following set of equations as our basic model,

$$\begin{aligned}
\frac{dT}{dt} &= s - dT - kVT,\\
\frac{dT^*}{dt} &= kVT - \delta T^* - d_x E T^*,\\
\frac{dV}{dt} &= N\delta T^* - cV,\\
\frac{dE}{dt} &= pT^* - d_E E.
\end{aligned} \tag{3}$$

The initial conditions are $T(0)=T_0$, $T^*(0)=0$, $V(0)=V_0$, $E(0)=0$. Here T, T^*, V and $E \in \mathbb{R}^+$ and all parameters are in $\mathbb{R}^+$. The constant s represents a source of healthy cells and d is their

death rate. k is the infectivity rate, δ is the death rate of the infected cells and d_x represents the effectiveness of the immune response. N is the number of virus particles produced per infected cell and c is the viral clearance rate.

The inclusion of the term $d_x ET^*$, allows for the removal of productively infected T cells due to a cell mediated immune response. Stafford et al. [1] provided for this response by using a complicated and seemingly arbitrary exponential function for effector cell killing where instead of the $-\delta T^* - d_x ET^*$ term they have $\delta = \delta_0 + \delta_1(V)$ where $\delta_1(V) = 0$ for $t < t_1$ and $\delta_1(V) = f(t)V$ for $t > t_1$ and t_1 is the assumed start time of the $CD8^+$ response. They then defined

$$f(t) = \frac{\beta}{1 + \kappa e^{-(t-t_1)/\Delta T_1}} - \frac{\beta}{1 + \kappa e^{-(t-t_1)/\Delta T_2}} \tag{4}$$

and comment that other forms for f can provide similar responses (Stafford model 2). What we have done differently is allow for a general solution of E.

2.3. Analysis of effector model

In the absence of the virus, i.e., non-infected steady state, the T cell population has a steady state value of

$$\widehat{T} = \frac{s}{d} \tag{5}$$

with all other variables having a steady state value of zero. Eq. (3) has two steady state solutions, the non-infected steady state $(\widehat{T}, 0, 0, 0)$ and the infected steady state $(\overline{T}, \overline{T}^*, \overline{V}, \overline{E})$ where

$$\begin{aligned}
\overline{T}^* &= \frac{kN\delta d_E}{cd_x p}\overline{T} - \frac{\delta d_E}{d_x p} \rightarrow \overline{T} > \frac{c}{Nk}, \\
\overline{V} &= \frac{N\delta}{c}\left(\frac{kN\delta d_E}{cd_x p}\overline{T} - \frac{\delta d_E}{d_x p}\right) = \frac{N\delta}{c}\overline{T}^*, \\
\overline{E} &= \frac{p}{d_E}\left(\frac{kN\delta d_E}{cd_x p}\overline{T} - \frac{\delta d_E}{d_x p}\right) = \frac{p}{d_E}\overline{T}^*,
\end{aligned} \tag{6}$$

where $\overline{T}$ is

$$\overline{T} = \frac{\frac{kN\delta^2 d_E}{cd_x p} - d + \sqrt{\left(\frac{kN\delta^2 d_E}{cd_x p} - d\right)^2 + 4s\frac{k^2N^2\delta^2 d_E}{c^2 d_x p}}}{2\frac{k^2N^2\delta^2 d_E}{c^2 d_x p}} \tag{7}$$

and is easily shown to be positive for all parameters.

In order to study the stability of the steady states we linearize (3) about a given steady state, (denoted by a subscript 'ss'), to get

$$A = \begin{pmatrix} -d - kV_{ss} & 0 & -kT_{ss} & 0 \\ kV_{ss} & -\delta - d_x E_{ss} & kT_{ss} & -d_x T^*_{ss} \\ 0 & N\delta & -c & 0 \\ 0 & p & 0 & -d_E \end{pmatrix}.$$

The characteristic equation for the linearized system is

$$\lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4 = 0, \tag{8}$$

where

$$\begin{aligned}
a_1 &= c + \delta + d_x E_{ss} + d_E + d + kV_{ss},\\
a_2 &= (d_E k + kc + k\delta + kd_x E_{ss})V_{ss} + dd_E + \delta d_E + d\delta + dc + cd_E + d_x E_{ss} c + T^*_{ss} p d_x\\
&\quad + dd_x E_{ss} + \delta c - \delta k T_{ss} N + d_x E_{ss} d_E,\\
a_3 &= ((kd_x c + kd_x d_E)V_{ss} + dd_x c + cd_E d_x + dd_x d_E)E_{ss} + (-\delta k N d_E - \delta dkN)T_{ss}\\
&\quad + (\delta d_E k + d_E kc + pkd_x T^*_{ss} + \delta kc)V_{ss} + \delta cd_E + (pdd_x + pd_x c)T^*_{ss}\\
&\quad + \delta d_E d + \delta dc + cd_E d,\\
a_4 &= (kcd_E d_x E_{ss} + \delta d_E kc + T^*_{ss} d_x cpk)V_{ss}\\
&\quad + dcd_E d_x E_{ss} + cd_E \delta d - \delta dk T_{ss} N d_E + T^*_{ss} p d_x cd.
\end{aligned}$$

By the Routh–Hurwitz conditions, all eigenvalues of (8) have negative real part if and only if

$$a_1 > 0, \quad a_4 > 0, \quad B_1 \equiv a_1 a_2 - a_3 > 0 \quad \text{and} \quad B_1 a_3 - a_1 a_4 > 0. \tag{9}$$

Proposition 1. *The non-infected steady state is stable if and only if*

$$\frac{c}{Nk} > \frac{s}{d}. \tag{10}$$

Proof. The Jacobian matrix at the non-infected steady state is

$$A = \begin{pmatrix} -d & 0 & -\frac{ks}{d} & 0\\ 0 & -\delta & \frac{ks}{d} & 0\\ 0 & N\delta & -c & 0\\ 0 & p & 0 & -d_E \end{pmatrix},$$

which produces the eigenvalues

$$\begin{aligned}
&\lambda_1 = -d, \quad \lambda_2 = -d_E,\\
&\lambda_{3,4} = \frac{-(c+\delta) \pm \sqrt{(c-\delta)^2 + \frac{4ksN\delta}{d}}}{2}.
\end{aligned} \tag{11}$$

Hence it is easily seen that all eigenvalues are real and negative given (10) is satisfied. Under this assumption the non-infected steady-state is locally stable. □

Proposition 2. *The infected steady state (6) is stable if and only if the Routh–Hurwitz inequalities (9) evaluated at this steady state are satisfied.*

Proof. From (6) we have that the infected steady-state exists if and only if

$$\frac{Nk}{c}\overline{T} > 1,$$

which will make $\overline{T}^*, \overline{V}, \overline{E}$ positive.

Let $y = \frac{Nk}{c}\overline{T} - 1$. We will show that a_1, a_4, $a_1a_2 - a_3$ are positive when y is positive. By substituting equalities (6) into the expressions of a_1, a_2, a_3 and a_4 we obtain

$$\begin{aligned}
a_1 &= y\left(\delta + \frac{kN\delta^2 d_E}{cd_x p}\right) + c + \delta + d + d_E,\\
a_2 &= y\left(\frac{k^2N^2\delta^3 d_E}{c^2 p d_x}\overline{T} + \frac{kN\delta^2 d_E}{pd_x} + \frac{kN\delta^2 d_E^2}{cpd_x} + \delta d_E\right)\\
&\quad + \frac{kN\delta d}{c}\overline{T} + \frac{kN\delta d_E}{c}\overline{T} + dd_E + cd_E + cd,\\
a_3 &= y^2\frac{kN\delta^3 d_E^2}{cpd_x} + y\left(\frac{k^2N^2\delta^3 d_E}{cpd_x}\overline{T} + \frac{k^2N^2\delta^3 d_E^2}{c^2 pd_x}\overline{T} + \frac{kN\delta^2 d_E^2}{pd_x} + \delta d_E c + \delta d d_E\right)\\
&\quad + cd_E d + \frac{kN\delta d d_E}{c}\overline{T},\\
a_4 &= y^2\frac{kN\delta^3 d_E^2}{pd_x} + y\left(\frac{k^2N^2\delta^3 d_E^2}{cpd_x}\overline{T} + c\delta d d_E\right).
\end{aligned}$$

We now have a_1, a_2, a_3 and a_4 as polynomials in y with positive coefficients, so they are positive for positive values of y. The next step is to compute $B_1 = a_1a_2 - a_3$ which is

$$\begin{aligned}
B_1 &= y^2\left(\frac{kN\delta^3 d_E^2}{cpd_x} + \frac{kN\delta^3 d_E}{pd_x} + \frac{k^2N^2\delta^4 d_E^2}{cp^2 d_x^2} + \frac{k^2N^2\delta^4 d_E^3}{c^2p^2d_x^2} + \frac{k^2N^2\delta^4 d_E}{c^2pd_x}\overline{T} + \frac{k^3N^3\delta^5 d_E^2}{c^3p^2d_x^2}\overline{T} + \delta^2 d_E\right)\\
&\quad + y\left(\delta cd + \delta d d_E + \delta d_E c + \delta^2 d_E + \delta d_E^2 + \frac{k^2N^2\delta^3 d_E^2}{c^2pd_x}\overline{T} + \frac{k^2N^2\delta^4 d_E}{c^2pd_x}\overline{T} + \frac{2k^2N^2\delta^3 d_E d}{c^2pd_x}\overline{T}\right.\\
&\quad + \frac{kN\delta^2 d}{c}\overline{T} + \frac{kN\delta^2 d_E}{c}\overline{T} + \frac{2kN\delta^2 d_E^2}{pd_x} + \frac{kN\delta^2 d_E^3}{cpd_x} + \frac{kN\delta^3 d_E}{pd_x} + \frac{kN\delta^2 d_E^3}{cpd_x} + \frac{ckN\delta^2 d_E}{pd_x}\\
&\quad \left. + \frac{2kN\delta^2 d_E d}{pd_x} + \frac{2kN\delta^2 d_E^2 d}{cpd_x}\right) + \overline{T}\left(\frac{kN\delta d_E^2}{c} + \frac{kN\delta^2 d_E}{c} + \frac{kN\delta d_E d}{c} + \frac{kN\delta^2 d}{c} + kNd\delta + kNd_E\delta + \frac{kN\delta d^2}{c}\right)\\
&\quad + cd_E^2 + cd^2 + dd_E^2 + c^2 d + c^2 d_E + c\delta d + d^2 d_E + 2cd_E d + \delta d d_E + \delta d_E c.
\end{aligned}$$

Here B_1 is a second degree polynomial in y with positive coefficients. Hence B_1 is always positive on the positive axis. The proof of the last relation $B_1a_3 - a_1a_2 > 0$ is tedious and can be verified using maple to be true when (10) is violated. □

We will show later that the stability of the infected steady state can bifurcate into a periodic orbit when we include a time delay.

3. Model with time delays

In this section we introduce a time delay in the model (3) by assuming that the immune response at time t is generated by the infection of a cell T^* at time $t-\tau$, where τ is constant. The model (Delay model 1) then becomes

$$\begin{aligned}
\frac{dT}{dt} &= s - dT - kVT,\\
\frac{dT^*}{dt} &= kVT - \delta T^* - d_x E T^*,\\
\frac{dV}{dt} &= N\delta T^* - cV,\\
\frac{dE}{dt} &= pT^*(t-\tau) - d_E E,
\end{aligned} \tag{12}$$

where the term $T^*(t-\tau)$ allows for a time delay between the moment of infection and the recognition of the infected cells by the cytotoxic $CD8^+$ T cells. The initial values are

$$T(0) = T_0, \quad T^*(\theta) = 0, \quad V(0) = V_0, \quad E(0) = 0, \ \ \theta \in [-\tau, 0].$$

3.1. Analysis of Delay model 1

We have again two steady-states, the non-infected steady-state $(\widehat{T}, 0, 0, 0)$ and the infected steady state $(\overline{T}, \overline{T}^*, \overline{V}, \overline{E})$ where $\overline{T}, \overline{T}^*, \overline{V}, \overline{E}$ are given by (6). We want to see if the delay changes the stability of the steady states. The linearized system in vector form is

$$\frac{dY}{dt} = BY(t) + CY(t-\tau),$$

where B and C are defined by

$$B = \begin{pmatrix} -d - kV_{ss} & 0 & -kT_{ss} & 0\\ kV_{ss} & -\delta - d_x E_{ss} & kT_{ss} & -d_x T^*_{ss}\\ 0 & N\delta & -c & 0\\ 0 & 0 & 0 & -d_E \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & p & 0 & 0 \end{pmatrix}$$

and

$$Y(t) = \begin{pmatrix} T(t)\\ T^*(t)\\ V(t)\\ E(t) \end{pmatrix}.$$

The characteristic equation for (12) is then given by

$$\det(\lambda I_4 - B - e^{-\lambda\tau} C) = 0,$$

that is,

$$H(\lambda) = \lambda^4 + b_1\lambda^3 + b_2\lambda^2 + b_3\lambda + b_4 + b_5\lambda^2 e^{-\lambda\tau} + b_6\lambda e^{-\lambda\tau} + b_7 e^{-\lambda\tau} = 0, \tag{13}$$

where

$$\begin{aligned}
b_1 &= a_1,\\
b_2 &= a_2 - pd_x T^*_{ss},\\
b_3 &= a_3 - (c + kV_{ss} + d)pd_x T^*_{ss},\\
b_4 &= a_4 - (kV_{ss} + d)pd_x cT^*_{ss},\\
b_5 &= d_x pT^*_{ss},\\
b_6 &= (kV_{ss} + d + c)d_x pT^*_{ss},\\
b_7 &= (d + kV_{ss})cd_x pT^*_{ss},
\end{aligned}$$

where a_1, a_2, a_3, a_4 are the coefficients of the characteristic polynomial for the effector model, i.e., $\tau = 0$.

It is known that the infected steady state is stable if all the eigenvalues given by (13) have negative real parts. We know that (13) with the condition $\tau = 0$ has all roots in the left half plane if (9) is satisfied. By Rouché's Theorem the transcendental equation has roots with positive real parts, given it had negative roots for $\tau = 0$, if and only if it has purely imaginary roots. Our goal is to find under what conditions we will have no such roots. Recent results have shown that this calculation can be made by applying sturm sequences (see [18]). Let $\lambda = \mu(\tau) + \mathrm{i}\nu(\tau)$ be the root of (13), with μ and $\nu \in \mathbb{R}$. If we substitute into (13) we get

$$\begin{aligned}
H(\mu,\nu) = {} & [\mu^4 + \nu^4 - 6\mu^2\nu^2 + b_1(\mu^3 - 3\mu\nu^2) + b_2(\mu^2 - \nu^2) + b_3\mu + b_4 + b_5 e^{-\mu\tau}((\mu^2 - \nu^2)\cos(\nu\tau)\\
& + 2\mu\nu\sin(\nu\tau)) + b_6 e^{-\mu\tau}(\mu\cos(\nu\tau) + \nu\sin(\nu\tau)) + b_7 e^{-\mu\tau}\cos(\nu\tau)]\\
& + \mathrm{i}[4\mu^3\nu - 4\mu\nu^3 + b_1(3\mu^2\nu - \nu^3) + 2b_2\mu\nu + b_3\nu + b_5 e^{-\mu\tau}(2\mu\nu\cos(\nu\tau) - (\mu^2 - \nu^2)\sin(\nu\tau))\\
& + b_6 e^{-\mu\tau}(\nu\cos(\nu\tau) - \mu\sin(\nu\tau)) - b_7 e^{-\mu\tau}\sin(\nu\tau)].
\end{aligned}$$

If we can find a root of (13) of the form $\lambda = \mathrm{i}\nu$, then the infected steady state losses its stability. This will happen if and only if

$$\begin{aligned}
H(0,\nu) = {} & [\nu^4 - b_2\nu^2 + b_4 - b_5\nu^2\cos(\nu\tau) + b_6\nu\sin(\nu\tau) + b_7\cos(\nu\tau)]\\
& + \mathrm{i}[-b_1\nu^3 + b_3\nu + b_5\nu^2\sin(\nu\tau) + b_6\nu\cos(\nu\tau) - b_7\sin(\nu\tau)] = 0
\end{aligned} \tag{14}$$

for some $\nu \in \mathbb{R}$. It is clear that we have a solution ν of (14) if both the real and imaginary parts of the equation equal zero. In other words,

$$\begin{aligned}
\nu^4 - b_2\nu^2 + b_4 &= b_5\nu^2\cos(\nu\tau) - b_6\nu\sin(\nu\tau) - b_7\cos(\nu\tau)\\
-b_1\nu^3 + b_3\nu &= -b_5\nu^2\sin(\nu\tau) - b_6\nu\cos(\nu\tau) + b_7\sin(\nu\tau).
\end{aligned}$$

If we add the square of both equations we obtain

$$\nu^8 + \alpha\nu^6 + \beta\nu^4 + \gamma\nu^2 + \omega = 0, \tag{15}$$

where $\alpha = b_1^2 - 2b_2$, $\beta = b_2^2 + 2b_4 - 2b_1b_3 - b_5^2$, $\gamma = -2b_2b_4 + b_3^2 - b_6^2 + 2b_5b_7$, and $\omega = b_4^2 - b_7^2$. If we substitute $w = \nu^2$ into (15) we obtain a fourth degree polynomial in w

$$w^4 + \alpha w^3 + \beta w^2 + \gamma w + \omega = 0. \tag{16}$$

If all the roots of the polynomial (16) are negative or imaginary then (15) will have no roots since v is real. Hence, we have no purely imaginary roots for H, and the infected steady state will be stable in the delay case.

If we reconsider (16), a fourth degree polynomial with real coefficients, we can apply the Routh–Hurwitz conditions to show when all the roots lie in the left half plane. Hence, we have the following result.

Proposition 3. *Under the following assumptions*:

$$\alpha > 0, \quad \omega > 0, \quad \alpha\beta - \gamma > 0, \quad \alpha\beta\gamma - \alpha\omega - \gamma^2 > 0.$$

Eq. (16) has no positive real roots.

4. Logistic model for target cells

In our Delay model 1 (12) we have assumed a constant source term for target cells and an exponential death rate, however, it is possible to consider other behaviors such as logistic growth [2]. Assuming target cells can grow at a rate, α day^{-1} and that this growth is limited by a carrying capacity, $T_{\max}$ cells, we have

$$\begin{aligned}
\frac{dT}{dt} &= \alpha T\left(1 - \frac{T + T^*}{T_{\max}}\right) - kVT,\\
\frac{dT^*}{dt} &= kVT - \delta T^* - d_x E T^*,\\
\frac{dV}{dt} &= N\delta T^* - cV,\\
\frac{dE}{dt} &= pT^* - d_E E
\end{aligned} \tag{17}$$

with initial conditions $T(0) = T_0$, $T^*(0) = 0$, $V(0) = V_0$, $E(0) = 0$. $T_{\max}$ is the total number of target cells in the plasma.

4.1. Analysis of logistic model

In the absence of the virus the T-cell population has a steady state value of $\widehat{T} = T_{\max}$. Now if we consider an infection modeled by the system (17) we will have two steady state solutions, the non-infected steady state $(\widehat{T}, 0, 0, 0)$ and the infected steady state $(\widehat{T}, \widehat{T}^*, \widehat{V}, \widehat{E})$ where

$$\begin{aligned}
\widehat{T}^* &= \frac{kN\delta d_E}{cd_x p}\widehat{T} - \frac{\delta d_E}{d_x p},\\
\widehat{V} &= \frac{N\delta}{c}\left(\frac{kN\delta d_E}{cd_x p}\widehat{T} - \frac{\delta d_E}{d_x p}\right),
\end{aligned}$$

$$\widehat{E} = \frac{p}{d_E}\left(\frac{kN\delta d_E}{cd_x p}\widehat{T} - \frac{\delta d_E}{d_x p}\right),$$

$$\widehat{T} = \frac{\alpha + \dfrac{kN\delta^2 d_E}{cd_x p} + \dfrac{\alpha d_E \delta}{d_x p T_{\max}}}{\dfrac{\alpha}{T_{\max}} + \dfrac{\alpha N\delta k d_E}{cd_x p T_{\max}} + \dfrac{k^2 N^2 \delta^2 d_E}{c^2 d_x p}}.$$

We linearize our system (where the subscript 'ss' is referring to any given steady state) to study stability and find the Jacobian matrix to be

$$A = \begin{pmatrix} \alpha - \dfrac{\alpha T^*_{ss}}{T_{\max}} - \dfrac{2\alpha T_{ss}}{T_{\max}} - kV_{ss} & -\dfrac{\alpha T_{ss}}{T_{\max}} & -kT_{ss} & 0 \\ kV_{ss} & -\delta - d_x E_{ss} & kT_{ss} & -d_x T^*_{ss} \\ 0 & N\delta & -c & 0 \\ 0 & p & 0 & -d_E \end{pmatrix}. \tag{18}$$

The corresponding characteristic equation for the linearized system is

$$\lambda^4 + \hat{a}_1\lambda^3 + \hat{a}_2\lambda^2 + \hat{a}_3\lambda + \hat{a}_4 = 0, \tag{19}$$

where

$$\hat{a}_1 = -\alpha + d_E + \delta + c + kV_{ss} + d_x E_{ss} + \frac{\alpha T^*_{ss}}{T_{\max}} + \frac{2\alpha T_{ss}}{T_{\max}},$$

$$\begin{aligned}\hat{a}_2 = {} & \left(\frac{\alpha d_x}{T_{\max}}T^*_{ss} + kd_x V_{ss} + \frac{2\alpha d_x}{T_{\max}}T_{ss} + d_x(c + d_E - \alpha)\right)E_{ss} + \frac{(\alpha c + \alpha\delta + T_{\max} p d_x + \alpha d_E)}{T_{\max}}T^*_{ss} \\ & + \left(\frac{k\alpha}{T_{\max}}T_{ss} + kc + kd_E + k\delta\right)V_{ss} + \left(\frac{2\alpha(c + d_E + \delta)}{T_{\max}} - N\delta k\right)T_{ss} + (-\alpha\delta + \delta d_E + cd_E - \alpha d_E + \delta c - \alpha c),\end{aligned}$$

$$\begin{aligned}\hat{a}_3 = {} & \left((kd_x c + kd_x d_E)V_{ss} + (d_x c d_E - \alpha d_x c - \alpha d_x d_E) + \frac{\alpha d_x c + \alpha d_x d_E}{T_{\max}}T^*_{ss} + \frac{2\alpha d_x c + 2\alpha d_x d_E}{T_{\max}}T_{ss}\right)E_{ss} \\ & + \left(\frac{(kd_E\alpha + k\alpha c)}{T_{\max}}T_{ss} + kpd_x T^*_{ss} + (k\delta c + kcd_E + k\delta d_E)\right)V_{ss} + (-\alpha\delta d_E + \delta c d_E - \alpha c d_E - \alpha\delta c) \\ & + \frac{\alpha p d_x}{T_{\max}}T^{*\,2}_{ss} + \left(\frac{(2\alpha p d_x - k\alpha N\delta)}{T_{\max}}T_{ss} + \frac{(\alpha c d_E + \alpha\delta c + T_{\max} p d_x c - \alpha T_{\max} p d_x + \alpha\delta d_E)}{T_{\max}}\right)T^*_{ss} \\ & + \frac{\alpha T_{\max} N\delta k + 2\alpha c d_E + 2\alpha\delta d_E - T_{\max} N\delta k d_E + 2\alpha\delta c}{T_{\max}}T_{ss} - \frac{2\alpha N\delta k}{T_{\max}}T^2_{ss},\end{aligned}$$

$$\begin{aligned}\hat{a}_4 = {} & \left(\frac{\alpha c d_x d_E}{T_{\max}}T^*_{ss} + kd_x c d_E V_{ss} - \alpha c d_x d_E + \frac{2\alpha d_x c d_E T_{ss}}{T_{\max}}\right)E_{ss} \\ & + \left(kpd_x cT^*_{ss} + k\delta c d_E + \frac{kcd_E\alpha T_{ss}}{T_{\max}}\right)V_{ss} \\ & + \alpha N\delta k d_E T_{ss} - \alpha\delta c d_E + \frac{2\alpha\delta c d_E}{T_{\max}}T_{ss} - \frac{2\alpha N\delta k d_E}{T_{\max}}T^2_{ss} \\ & + \frac{\alpha p d_x c}{T_{\max}}T^{*\,2}_{ss} + \frac{\alpha\delta c d_E + 2\alpha c p d_x T_{ss} - \alpha\delta N k d_E T_{ss} - \alpha p c d_x T_{\max}}{T_{\max}}T^*_{ss}.\end{aligned}$$

Proposition 4. *The non-infected steady state is stable if and only if*

$$\frac{c}{Nk} > T_{\max}. \tag{20}$$

Proof. The Jacobian matrix at the non-infected steady state $(T_{\max}, 0, 0, 0)$ is

$$A = \begin{pmatrix} -\alpha & -\alpha & -kT_{\max} & 0 \\ 0 & -\delta & kT_{\max} & 0 \\ 0 & N\delta & -c & 0 \\ 0 & p & 0 & -d_E \end{pmatrix},$$

where the characteristic equation

$$(\lambda+\alpha)(\lambda+d_E)(\lambda^2+(\delta+c)\lambda+\delta c - NkT_{\max}\delta) = 0$$

has solutions

$$\lambda_1 = -\alpha, \quad \lambda_2 = -d_E,$$
$$\lambda_{3,4} = \frac{-(c+\delta) \pm \sqrt{(c+\delta)^2 - 4\delta c + 4NkT_{\max}\delta}}{2}.$$

If $\frac{c}{Nk} > T_{\max}$ then all eigenvalues are negative so the non-infected steady state is stable. This means that the virus will eventually go extinct. When $\frac{c}{Nk} < T_{\max}$ the non-infected steady state is unstable so the virus will grow without any control. We see (20) as implying that the rate of clearance of the virus is bigger than the rate of infection. □

Proposition 5. *The infected steady state is stable if and only if the following inequalities are satisfied,* $\hat{a}_1 > 0,\ \hat{a}_4 > 0,\ \hat{a}_1\hat{a}_2 - \hat{a}_3 > 0,\ \hat{a}_1(\hat{a}_2\hat{a}_3 - \hat{a}_4) - \hat{a}_3^2 > 0.$

Next we consider the logistic model with a time delay (Delay model 2) as before and study

$$\begin{aligned} \frac{dT}{dt} &= \alpha T\left(1 - \frac{T+T^*}{T_{\max}}\right) - kVT, \\ \frac{dT^*}{dt} &= kVT - \delta T^* - d_x E T^*, \\ \frac{dV}{dt} &= N\delta T^* - cV, \\ \frac{dE}{dt} &= pT^*(t-\tau) - d_E E, \end{aligned} \tag{21}$$

where τ is the time between the infection of the target cells and the initiation of the immune response. Using the same arguments as in Section 3 we find two steady states, the non-infected state $(\widehat{T}, 0, 0, 0)$ and the infected state $(\widehat{T}, \widehat{T}^*, \widehat{V}, \widehat{E})$ providing the characteristic equation

$$\lambda^4 + b_1\lambda^3 + b_2\lambda^2 + b_3\lambda + b_4 + b_5\lambda^2 e^{-\lambda\tau} + b_6\lambda e^{-\lambda\tau} + b_7 e^{-\lambda\tau} = 0, \tag{22}$$

where

$$\begin{aligned}
b_1 &= a_1,\\
b_2 &= a_2 - pd_x T^*_{ss},\\
b_3 &= a_3 - 2\frac{\alpha pd_x T^*_{ss} T_{ss}}{T_{max}} - pd_x T^*_{ss} c - kV_{ss} pd_x T^*_{ss} + \alpha pd_x T^*_{ss} - \frac{\alpha pd_x}{T_{max}} T^{*\,2}_{ss},\\
b_4 &= a_4 + \alpha pd_x T^*_{ss} c - kV_{ss} pd_x T^*_{ss} c - 2\frac{\alpha T_{ss} pd_x T^*_{ss} c}{T_{max}} - \frac{\alpha pd_x c}{T_{max}} T^{*\,2}_{ss},\\
b_5 &= pd_x T^*_{ss},\\
b_6 &= \frac{2\alpha T_{ss} pd_x T^*_{ss}}{T_{max}} + \frac{\alpha T^{*\,2}_{ss} pd_x}{T_{max}} - \alpha pd_x T^*_{ss} + kV_{ss} pd_x T^*_{ss} + pd_x T^*_{ss} c,\\
b_7 &= \frac{2\alpha T_{ss} pd_x T^*_{ss} c}{T_{max}} - \alpha pd_x T^*_{ss} c + \frac{\alpha T^{*\,2}_{ss} pd_x c}{T_{max}} + kV_{ss} pd_x T^*_{ss} c.
\end{aligned}$$

5. Parameter estimations

We opted to implement a directed Monte Carlo to optimize 7 of the 10 unknown parameters of the Delay model 1 (12), and the 8 unknown parameters of the Delay model 2 (21). For (12) we consider the parameter estimates found in Stafford et al. for their model of primary infection and chose to fix s, d, and k to be the same as given in [1]. The infectivity rate, k, has been shown to be reliable using the steady state condition, $k = \frac{c}{NT_0}$. The parameters s and d relate to target cell dynamics and since we are only fitting viral load data, we chose to fix these as well. Hence, the focus of the parameter estimations is dealing with viral load kinetics. For review, we present the general model Stafford et al. used for primary infection (Stafford model 1) [1].

$$\begin{aligned}
\frac{dT}{dt} &= s - dT - kVT,\\
\frac{dT^*}{dt} &= kVT - \delta T^*,\\
\frac{dV}{dt} &= N\delta T^* - cV.
\end{aligned} \tag{23}$$

They considered T_0 to be 10 000 cells/ml, T^*_0 to be 0 cells/ml, V_0 to be 10^{-6} virus/ml, and c to be 3 day^{-1}. To estimate the remaining parameters they used data from ten patients and the method of least square estimation. We present as a reference only their median values for all ten patients.

The parameters given in Table 1 served as a platform to estimate values for our unknown parameters using a hill climbing Monte Carlo algorithm. To optimize model parameters using this algorithm, one first defines a fitness function that given arbitrary parameters, returns how closely

Table 1
The median values for certain parameters for the 10 patients found in [1]

Parameter	d (day^{-1})	δ (day^{-1})	k (ml virion-day^{-1})	$(\pi/\delta)N$ (virus per cell)	$(\lambda)s$ (cells μl^{-1} day^{-1})
Summary of median values from data estimations in Stafford et al.					
Median	0.01	0.39	0.00000065	2241	0.14125
SD	0.0057	0.19	0.0000014	4339	0.08761

In parenthesis is the notation used in [1] and next to it is our notation. It should be noted that in Stafford et al. [1] they set $\pi = N\delta$, and gave estimates for π. Hence, the values for N given above are determined by π/δ.

the model using these parameters fits observations. Each parameter is then selected at random from a pre-defined interval and the fitness of that parameter set is evaluated. Many parameter sets are then tested with the hope of finding the set that allows the model to most closely match observations (see [19] for a complete sensitivity analysis of the model parameters).

This was done by first using the algorithm to search only within a range of values for each parameter that we determined to be reasonable, and once it found an acceptable fit, only select parameter candidates from within a particular range of the value of each parameter used in the current 'best model'. That is, once a parameter set had been discovered with reasonable fitness, the algorithm would only explore locally to those parameters to select its next candidate. By varying the size of the 'window' from which the next parameters could be chosen, we could both search the entire parameter space and local portions to maximize fitness by making very small changes in the parameters. This algorithm also has the advantage of avoiding stagnation in local maxima because it is seeded by the results of a global search, and because one can increase the window width to find regional fits better than the local maximum.

The fitness, f, of a particular parameter set was chosen as follows:

$$1/f = \left(\sum_{i=1}^{n} (\log VTiter(t_i) - \log V(t_i))^2 + (n-2)(\log VTiter_{\max} - \log V_{\max})^2 \right)^{\frac{1}{2}}, \tag{24}$$

where $VTiter(t_i)$ is the viral load in each patient at time t_i when a sample was taken, $V(t_i)$ is the viral load predicted by the model at the time t_i, given particular parameters, and calculated using a modified Runga–Kutta method with step size $\Delta t = 0.01$, $T_0 = 10$ cells/μl, $T_0^* = 0$ cells/μl, $E_0 = 0$ cells/μl, $V_0 = 10^{-9}$ virus/μl. $\text{ViralTiter}_{\max}$ is the maximum viral titer observed in the patient, $V_{\max}$ is the viral load predicted by the model for the time the patient's viral titer was observed to be maximal, and n is the number of observations of viral titer made before day 120.

We amplified the importance of the maximal viral observation in determining the fitness because in many of the patients, peak viremia was represented with very few data points, relative to the number of points where viral load was reduced from its peak. Given this distribution of points, the model might find solutions in which all data points but the maximal one were well predicted, thus minimizing a function that did not over-represent the importance of the maximal observation.

We first iterated the Monte Carlo algorithm (5000 times for the Delay model 1 (12) and 15 000 times for the Delay model 2 (21)) with the following parameter ranges: $\delta \in (0.3, 7)$, $c \in (2, 50)$,

$d_x \in (0,1)$, $p \in (0,2)$, $d_E \in (0,5)$, $\tau \in (0,28)$, and $k \in (0,2)$ (in the Delay model 2 only). In the Delay model 1, k, N, and s, and d were determined by Stafford's median or patient-specific parameter estimates. In the Delay model 2, $\alpha = 0.1$ and $T_{\max} = 1000$.

Next, the Monte Carlo chose parameters out of moving windows centered on the current 'best parameters' with width equal to $0.01x$ that of the initial parameter selection range. In both models 10000 iterations were performed with the sliding window. In most cases, this search regimen was sufficient to find a very close fit of the data. In more stubborn cases, we varied the number of search iterations and window widths until an acceptable fit was found. Once we had determined optimal parameters for each patient and in each model, we numerically investigated the sensitivity of both models to perturbations in each parameter. This was done by collecting the median values of each parameter for all the patients, and running the model with these parameters. The viral loads from this run were considered observations, and the viral loads from time $12, 16, \ldots, 120$ were used to calculate fitness as described above.

5.1. Convergence, significance and sensitivity of our estimates

It is impossible to guarantee that the sequence generated by our optimization method converges to the global maximum. To increase the confidence in our results, we repeatedly (randomly) seeded initial estimates for the optimization algorithm. In over 90% of the seedings, the algorithm converged to approximately the same value, thus supporting the notion that we have identified (at the very least) a local minimum.

A primary emphasis of this paper is to address how the introduction of additional biological phenomena into an HIV model impacts its predictions. One way to provide evidence for or against a more complex model is to consider the statistical issues related to how well a given model describes the data. We assume the following statistical model for the measurement error in each viral load observation

$$\log(VTiter(t_i)) = \log(V(t_i)) + \varepsilon_i; \quad i = 1, \ldots, n, \tag{25}$$

where the errors ε_i are independent and identically distributed (i.i.d.) random variables. The observations are made at the predetermined timepoints t_i, where $i = 1, \ldots, n$. Since Eq. (12) with parameters $[d_x, p, d_E, \tau]$ all set to zero is equivalent to Eq. (23), the models we consider are nested (where one model is a simplification of another). Within a nested structure, there are well-known inference methods for dealing with simple linear models, however, our models consist of non-linear differential equations without closed form solutions. Following the technique developed in [20], we define our least squares cost functional:

$$J_{n,m}(x) = \sum_{i=1}^{n} (\log(VTiter_{t_i}(m)) - \log(V_{t_i}(x)))^2, \tag{26}$$

where $x = (c, \delta, k, d_x, p, d_E, \tau)$ is the vector of candidate parameters, n is the number of data points, and m is the patient identification number.

A significant concern in [20] is the convergence of the numerical simulation scheme. Accordingly, they define a solution to their model equation as x and an approximate solution as x^N, for index of approximation N. As $N \to \infty$, they assume the chosen numerical scheme converges

uniformly to the actual solution. Subsequently, they prove the convergence of the optima of the approximate cost functional to the optima of the original functional (in distribution). Since our models are relatively simple and non-stiff in the chosen parameter regions, we assume that our modified Runga–Kutta method is sufficiently accurate to ensure that for all t of interest, $\|V_t^N - V_t\|$ (and by extension the difference between $J_{n,m}^N$ and $J_{n,m}$) is negligible.

We next define (as in [20]), the test statistic

$$U_n = n\frac{J_{n,m}(x^*_{\text{Stafford}}) - J_{n,m}(x^*_{\text{model 1}})}{J_{n,m}(x^*_{\text{model 1}})} \tag{27}$$

for x^*_{Stafford}, $x^*_{\text{model 1}}$ vectors of parameters which are the optimal least squares fit values for the parameters in Eqs. (23) and (12), respectively. By Theorem 4.6 in [20], we know that this statistic converges asymptotically to a $\chi^2(4)$ distribution as $n \to \infty$. We can, therefore, use a null hypothesis framework to identify statistically significant reductions in the least squares cost functional.

For patient 1, we find the least squares (LS) for Stafford's model to be 6.1 and the LS for our model to be 2.37. If we then consider the *null hypothesis* to be Stafford's fit given the optimal parameters we see that (27), i.e., a $\chi^2(4)$ test, allows us to reject this hypothesis at a useful confidence level since the statistic takes (assuming the null hypothesis that Stafford's model is correct) the value of

$$U_{17} = 26.75. \tag{28}$$

Recall that the data used in the analysis of Stafford's model is given in Table 1 of Stafford et al.

Hence, this suggests for patient 1, that the parameters obtained by our model are of statistical importance. Hence, the improvement in the fit in our model by adding the effector cell population it is not simply due to the increase of the degree of freedom in the model. We find similar conclusions follow in patient 3, 8 and 10. However the improvement in the fit in patients 2, 4, 5, 7, and 9 it is not statistically significant.

We then ran through a jackknifing analysis for each of the patients. In particular, for each execution of the parameter optimization, 5000 chains were initialized within a range of parameters we felt would encompass all likely values of the parameters. Of these, the one with best fitness, was used to initiate a Markov Chain, with a random step size chosen from a uniform distribution across the nearest 1% of the initial parameter ranges to the current parameter set in the chain. If the parameter set after the step had a higher fitness it was accepted with probability 1, otherwise, the chain stayed at the original parameter set. When the algorithm had considered at least 2000 sequential candidate parameter sets without a fitness improvement, we considered it to have converged and selected those parameters as our estimate.

For patients 3, 6, 7, 8 and 10 we left out 37% of the data points as these patients contained the most number of data points. For patients 2, 4, 5, and 9, we randomly removed 2 of the data points so that we could maintain at least 4 or 5 points for the fitting. In patient 1 we were not able to remove any points. The results of our jackknifing were much less conclusive. For two of the main parameters we are examining, i.e., the death rates of productively infected T cells, δ and virus, c, we found good consistency with the patient specific values as given in the next section, but both had a confidence interval that included both our values and the ones in Stafford. Hence, we could not infer any statistical significance between our results and those in Stafford [1]. For the produc-

tion and death rates of effector cells, p, and d_E respectively we found much more consistency as the jackknifing provided very low p-values when comparing the distributions of estimates between patients using a t-test. The parameter for the time lag for effector cells to become active, τ also had a high degree of consistency, i.e., low p-values but had relatively large standard errors which gives a wide confidence interval. However, the best fit values for τ between patients was very consistent, ranging from 19 to 32 days.

Hence we conclude that while our algorithm for fitting the patient data suggests we have found a new set of biologically relevant parameters, however, these estimates are not robust under all statistical analyses. This illustrates our main point of this paper; different models can provide widely varying estimates for parameters while predicting dynamics that fit the available data quite closely. Furthermore, the statistical frailty of some of these estimates makes it very hard to determine a favored model. That said, however, we have shown that our results are just as acceptable as others and can provide a drastically different interpretation of the biology; an interpretation that does need to be taken seriously.

6. Results and conclusions

We found two major differences between Stafford's model and our two models (12) and (21). First, we find in patients 3, 4, 5, 7, 8, and 9 (see Figs. 1 and 2) similar dynamical behavior between the Stafford model and our Delay model 1 (12) but with vastly different parameter estimations for certain parameters (we denote estimated values for this subgroup of patients by the subscript $e1$). For instance, we find the average estimated value for δ for these 6 patients to be $\delta_{e1} = 0.462$ day^{-1} compared to Stafford's $\delta_{e1} = 0.269$ day^{-1}. We estimated the time delay for these patients to be $\tau_{e1} = 8.29$ days. The most dramatic difference in parameters was seen in the death rate of virus, c. Stafford et al. assumed c was constant for all patients and set its value to $c = 3$ day^{-1}. We allowed the Monte Carlo simulations to fit this parameter and found $c_{e1} = 0.366$ day^{-1}. This is a difference of over 700%. This discrepancy in c suggests a major difference in the mechanism of pathogenesis. Assuming $c = 3$ day^{-1} is consistent with drug perturbation models [21–23], but if c were much smaller, it would suggest that viral particles are living longer during primary infection, i.e., about 2.73 days, while the productively infected T cells are living roughly half the time as predicted in Stafford et al. (see Table 2). Since both models fit the observed data closely, it is almost impossible to determine which is 'better', dramatically illustrating how any conclusions drawn are contingent on the structure of the model being considered. There are model selection methodologies on the statistical inference and Bortz and Nelson [24] have recently developed a methodology for selecting models based on their dynamical and statistical properties.

The second difference was seen in patients 1, 2, 6, and 10. It has been postulated that viral load levels in the plasma do not present any sustained oscillations after primary infection. This may be true simply because the data is often too sparse and sampled periodically. We suggest these oscillations are plausible since the immune response, specifically the CD8$^+$ T cells, can act in a predator prey type dynamic with the virus. However, to confirm this would require data to be sampled more frequently and in a somewhat random nature. Our Delay model 1 (12) predicts oscillations in the viral loads with periods between 30 and 60 days. This period is consistent with CD8$^+$ T cells

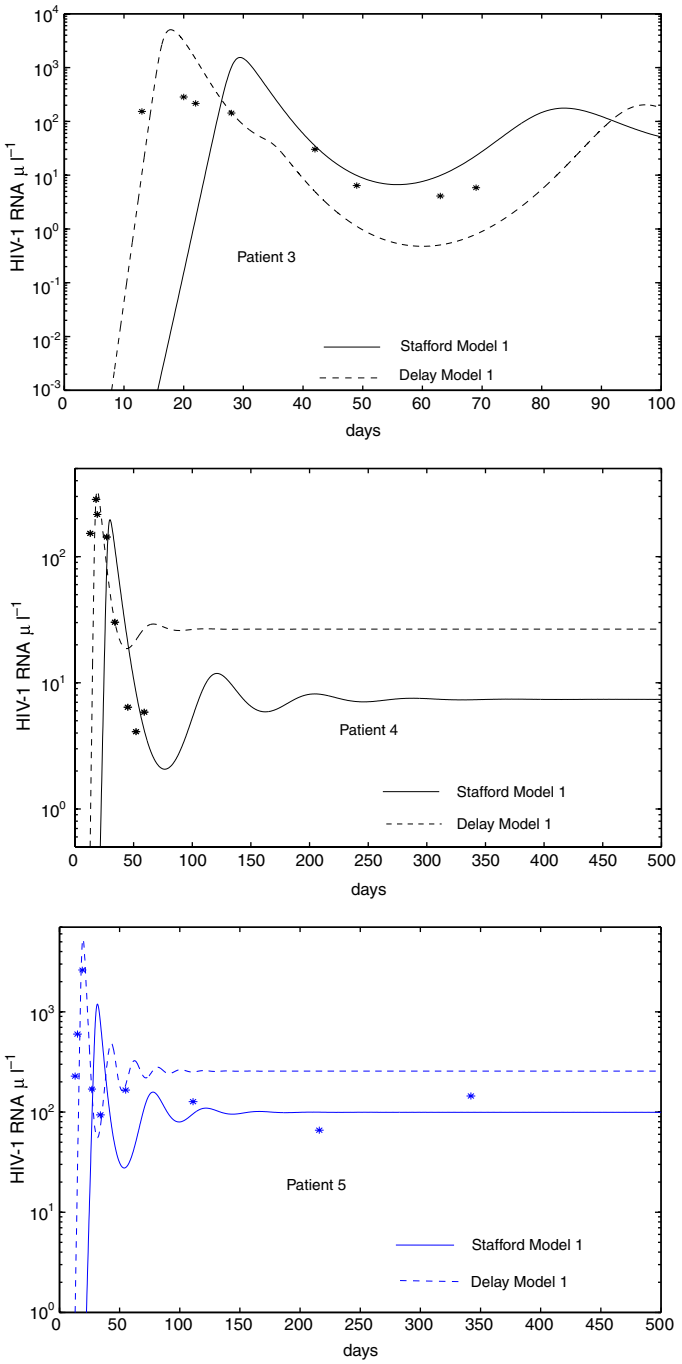

Fig. 1. The fits for patients 3, 4 and 5 using (12) (Delay model 1) and the Stafford model without the immune response (Stafford model 1) are comparable except that (12) seems to fit the earlier data points while the Stafford model seems to fit the later data points. In either case, the overall fitness is similar, however, there is a large variation in the estimated parameters, specifically for δ (see Section 6). All the parameters for each graph can be found in Table 2, with the initial conditions of $T(0) = 10\ \mu l^{-1}$, $T^*(0) = 0\ \mu l^{-1}$, $V(0) = 10^{-6}\ ml^{-1}$ and $E(0) = 0.1\ \mu l^{-1}$. Asterisks (*) are patient data.

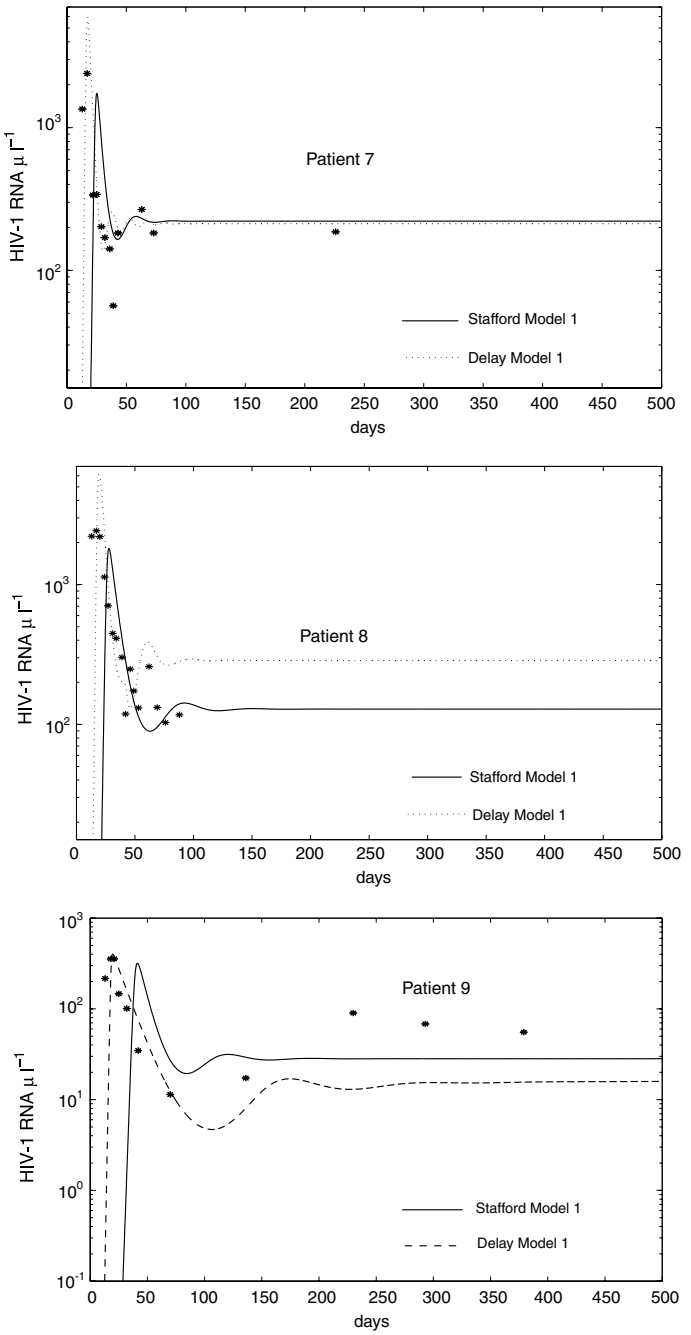

Fig. 2. The fits for patients 7, 8 and 9 using (12) (Delay model 1) and the Stafford model without the immune response (Stafford model 1) are again comparable as discussed in Fig. 1. See the caption in Fig. 1 for listed parameters. Asterisks (*) are patient data.

Table 2
Monte Carlo results for (12) where all parameters are estimated except for d, k, and s which we used their listed values in Stafford et al. [1]

Patient	δ	d_x	c	p	d_E	τ	N	d	k	s	Fitness	n	Adj. fitness
Monte Carlo results for patient data													
1	0.247	0.285	2.428	1.271	2.67	20.87	3969.04	0.013	0.00046	0.13	0.068	4	0.028
2	0.486	1.076	1.679	1.216	1.746	13.53	3703.49	0.02	0.00036	0.2	1.328	6	0.420
3	0.687	0.905	0.577	0.103	0.004	18.57	1397.94	0.0065	0.00064	0.065	1.570	9	0.393
4	0.612	1.667	0.246	3.912	8.087	0.004	160.26	0.0046	0.0048	0.046	1.398	8	0.374
5	0.541	0.623	0.537	1.768	0.403	3.3	1606.99	0.017	0.00063	0.17	0.886	7	0.256
6	0.711	0.72	0.699	2.288	1.490	19.14	1111.08	0.012	0.00075	0.12	2.154	11	0.482
7	0.345	1.595	0.372	2.947	0.471	2.51	2116.06	0.017	0.0008	0.17	0.694	11	0.155
8	0.588	0.388	0.387	1.258	3.059	24.96	1410.59	0.0085	0.00066	0.085	2.181	16	0.398
9	0.00019	1.444	0.081	1.675	0.0061	0.93	591851	0.006	0.0025	0.06	1.629	7	0.471
10	0.3	0.241	5.191	0.0051	1.798	27.98	23599.9	0.0043	0.00019	0.043	1.682	11	0.376
Median	0.514	0.812	0.557	1.473	1.618	16.05	1861.53	0.01025	0.00065	0.103	1.484		0.384
SD	0.225	0.536	1.572	1.193	2.34	10.63	185915	0.006	0.0014	0.0573	0.655		0.146

Confidence intervals are presented in Table 3.

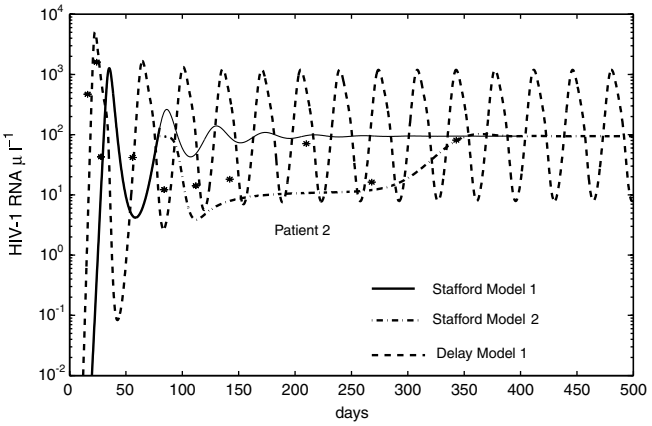

Fig. 3. Analysis for patient 2, comparing the models in Stafford et al. [1] and (12). The parameters for the Stafford model with (Stafford model 2) and without (Stafford model 1) an immune response are the same as in [1]. The Delay model 1 parameters are given in Table 2 for this patient. The Monte Carlo simulations verified our analysis that shows the existence of a Hopf bifurcation that allows for the model to predict sustained oscillations. We found oscillations with a period of about 45 days. The biggest difference between these models is the estimation of the parameters. For instance, we found in our model (12) a 79% reduction in the value for c, which suggests that during primary infection, the viral turnover is longer then during drug therapy. (However, if we compare this to the logistic model (21) we find over a 500% reduction in c when compared to $c = 3\ \text{day}^{-1}$ (see Fig. 6).) Asterisks (*) are patient data.

needing to adapt to newly evolved HIV-1 epitopes displayed on the surface of infected cells. It is also possible, as suggested in patient 2 (see Fig. 3), that there exists sustained oscillations, via a

Table 3
The 95% confidence intervals assuming the estimates are log-normally distributed for results given in Table 2

Patient		δ	d_x	c	p	d_E	τ	N
1	Lower	0.0226789	0.160276738	0.099839484	0.260005469	0.95010047	5.120048794	366.8748885
	Upper	1.348031651	0.96801883	24.20152985	2.453448009	6.929672732	62.59489331	22773.11784
2	Lower	0.236204182	0.127537921	1.456426047	0.589421222	0.106218752	5.260763448	2785.297951
	Upper	0.64625022	2.733636615	2.266810412	3.971569511	13.11810314	16.21598894	7620.532914
3	Lower	0.646922757	0.069259416	0.596930697	0.119174514	0.004944589	19.41844686	1015.203119
	Upper	0.945622429	5.737899322	0.929037465	15.56291339	5.708161666	63.83336503	1483.94903
4	Lower	0.275352203	0.038954183	0.203768322	0.040255426	0.008195895	0.000530818	135.6304597
	Upper	0.722552865	6.084658937	0.627359867	17.89462854	99.92980499	3651.242472	355.9075807
5	Lower	0.421197788	0.108687719	0.055676373	0.276838839	0.006517569	0.629862335	1302.772872
	Upper	0.66780674	2.502114282	1.663526698	4.559689107	50.92519049	47.3501406	2065.538923
6	Lower	0.618627043	0.39546588	0.647558041	0.6605023	0.370586396	9.79910631	995.1854994
	Upper	0.793821647	1.629274128	0.805892175	3.755358387	3.974179923	66.28099514	1277.020338
7	Lower	0.215508899	0.231343373	0.379419111	0.567989982	0.27761767	0.922829144	767.6616452
	Upper	0.950938082	2.785506687	0.636620521	4.447940205	6.422270402	180.2586609	3387.328389
8	Lower	0.130883905	0.063721435	0.150188955	0.521396896	0.005729389	0.863853909	689.2041245
	Upper	1.204289495	8.088723739	0.720121312	7.535534274	151.1571643	221.1404848	6341.486353
9	Lower	0.199842498	1.08060127	0.062986129	1.927227138	6.45266E−05	0.97020651	549.1158594
	Upper	0.200322267	1.889253583	0.080755748	3.911962959	0.0212772	1.581470339	550.4329246
10	Lower	0.135247436	0.007172789	5.238669787	0.314477336	0.948254932	4.301307489	27644.18074
	Upper	0.256835709	6.327643794	5.722874361	1.354280686	7.391867127	18.74529946	52496.35709

Confidence intervals for the parameters given in Table 2.

Table 4
Monte Carlo results for (21) where all parameters are estimated except for α and $T_{\max}$

Patient	δ	d_x	c	p	d_E	τ	N	k	α	$T_{\max}$	Fitness
Monte Carlo results for patient data											
1	0.508	0.235	13.799	1.575	3.273	0.677	1927.93	0.001124	0.1	1000	0.034
2	0.358	1.447	19.046	2.075	1.44	0.539	5028.74	0.000992	0.1	1000	1.339
3	0.166	0.318	2.853	0.857	3.5	0.009	5793.4	0.000567	0.1	1000	1.571
4	0.33	2.232	0.186	3.031	0.819	0.375	297.22	0.002493	0.1	1000	0.297
5	0.196	4.048	0.51	6.124	15.962	0.003	4431.37	0.000299	0.1	1000	0.842
6	0.102	0.763	2.529	0.621	5.166	0.0277	7780.57	0.000655	0.1	1000	1.164
7	1.14E−05	1.022	0.591	1.671	2.11	0.032	6.39E+07	0.00039	0.1	1000	0.537
8	0.0006	2.621	0.106	4.002	0.253	0.605	1.43E+06	0.000297	0.1	1000	0.71
9	0.021	1.270	0.517	0.408	3.138	0.043	5281.81	0.002206	0.1	1000	0.431
10	0.039	0.61	6.698	0.271	2.148	5.637	181457	0.000204	0.1	1000	0.749
Median	0.134	1.146	1.559	1.623	2.643	0.209	5537.61	0.00061	0.1	1000	0.7674
SD	0.176	1.197	6.618	1.858	4.512	1.723	20145215	0.000812	0	0	0.4789

Confidence intervals are presented in Table 5.

Hopf bifurcation, that implies the viral dynamics and the immune system are in a steady limit cycle. This sustained oscillation arguably fits the data better than the Stafford's model that allows

Table 5
The 95% confidence intervals assuming the estimates are log-normally distributed for results given in Table 4

Patient		δ	d_x	c	p	d_E	τ	N	
1	Lower	0.355926278	0.37702851	3.009281286	0.577063217	0.75718755	0.054254609	885.0862982	0.000808716
	Upper	1.107236216	1.184637443	55.42227009	2.399269234	9.045585368	4.056135149	2753.376546	0.001511673
2	Lower	0.128094369	0.543557027	10.6493692	1.178125104	0.310549966	0.124676494	1326.254336	0.000780159
	Upper	1.357206015	1.448264787	23.9349429	3.036814271	1.9540913	0.853590556	14052.1425	0.001200222
3	Lower	0.068707705	0.302853445	0.669431445	1.058533845	0.233500376	0.002162132	492.449679	0.000423118
	Upper	1.949438323	2.144297034	96.87491046	3.137505393	4.073558773	5.968913961	13972.23874	0.006949096
4	Lower	0.016977163	0.132255961	0.160247958	0.398792277	0.39060487	0.002389622	389.2911258	0.002043234
	Upper	0.25173926	2.902863421	2.791044798	4.322451946	25.76764482	2.547235282	5772.465441	0.00399994
5	Lower	0.011390414	0.734028369	0.391152095	1.111007602	4.661651903	0.000269476	1322.905442	0.000281053
	Upper	0.657644241	3.765393184	1.056756425	10.84105534	12.91344416	0.045092588	76379.97593	0.000330484
6	Lower	0.049436962	0.057711926	0.469614932	0.460493656	1.483131634	0.000823213	2861.206244	0.000344831
	Upper	0.276107527	3.132022799	6.871009494	4.111276677	9.638879843	3.129113844	15979.97769	0.001079469
7	Lower	4.21369E−06	0.677793968	0.417692617	1.11729361	0.384326674	0.001851063	69.32549911	0.00037794
	Upper	10.53004049	1.459692361	0.942442996	3.070994414	8.55633811	0.457018517	173244432.6	0.00044031
8	Lower	1.15049E−05	0.118998842	0.290142498	0.319610726	2.323238485	0.02233281	360187.1407	0.000330759
	Upper	0.002304357	2.923134007	1.887115323	4.964050869	5.403618194	0.745303254	72142898.37	0.000500491
9	Lower	0.000942735	0.390351303	0.179190962	0.458751288	2.484614502	0.006198279	1121.677211	0.001874126
	Upper	0.098067302	1.72273365	0.937811286	2.785938466	5.364394123	2.424903822	116681.8135	0.002471544
10	Lower	0.00930342	0.169269222	3.399139174	0.001207637	0.497764664	2.953410201	41273.23149	0.000120942
	Upper	0.17202424	1.473820091	18.58834758	26.99863384	14.00511861	13.57648783	763159.3491	0.000474163

Confidence intervals for the parameters given in Table 4.

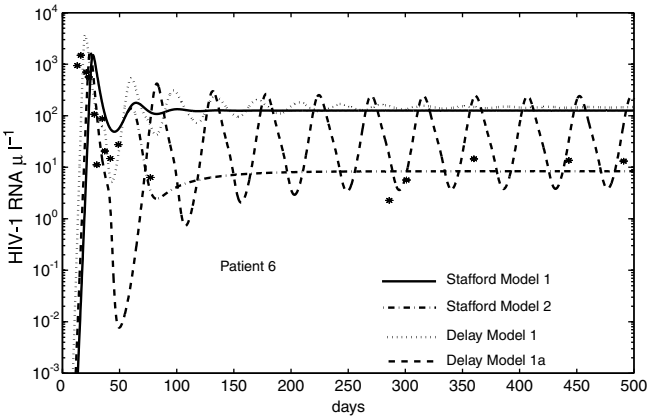

Fig. 4. Analysis for patient 6, comparing the models in Stafford et al. [1] and (12). The parameters for the Stafford model with (Stafford model 2) and without (Stafford model 1) an immune response are the same as in [1]. Our model's (12) (Delay model 1) parameters are given in Table 2 for patient 6. Fits of this data did not allow for the existence of sustained oscillations and matched observations during primary infection but not the long term data. Stafford's model without the immune response shows the same behavior and they corrected this fit by invoking an exponential immune response (Stafford model 2). Our model with the immune response did not fit well the long term data, but if we change the value for c from 0.69 day^{-1} (Delay model 1) to 1.39 day^{-1} (Delay model 1a), we find the sustained oscillations that provide a better fit to the data. Asterisks (*) are patient data.

for an immune response [1] (see Fig. 3) and supports our hypothesis of an adapting immune response (see Tables 3–5).

The analysis of patient 6's data (see Fig. 4) is interesting in that there are sampled data points early in the infection as well as a few long term data points, i.e., up to one year. Stafford's model without the immune response and our Delay model 1 using the estimated parameters from the Monte Carlo simulations fitted the early data well but not the long term data. Stafford was able to correct this by adding a non-autonomous term, i.e., a piecewise exponential term, to account for the immune response. Our Delay model 1 (12), which allowed the immune response to be a variable in the model, was also able to correct this problem in patient 2 but did not do such a good job in patient 6. However, if we adjusted the value for c from $c = 0.69$ day^{-1} to $c = 1.69$ day^{-1}, then we cross through a Hopf bifurcation which allowed for sustained oscillations and provided a closer fit then the Stafford model to the data (see Fig. 4).

A third comparison from patients 1 and 10 showed damped oscillations but no sustained oscillations for a realistic parameter set. The predicted behaviors of (12) and the Stafford model were again similar but again with changes to the parameter estimations (see Fig. 5). We found an average value for δ for these two patients to be $\delta_{e2} = 0.273$ while Stafford found $\delta_{e2} = 0.43$. Again suggesting the productively infected T cells are living longer during primary infection than during drug therapy or latency. While for some patients we found dramatic differences in the value for c, in these two patients we found a difference of only 20%, suggesting the virus particles are dying earlier in their life cycle (see Table 2), and that different viral dynamics may exist from one patient to another.

Finally, we examined how our predictions change when we allow for T cells to have a logistic growth (21). We chose the dynamics of the target cells to be the same as in [2]. We ran Monte

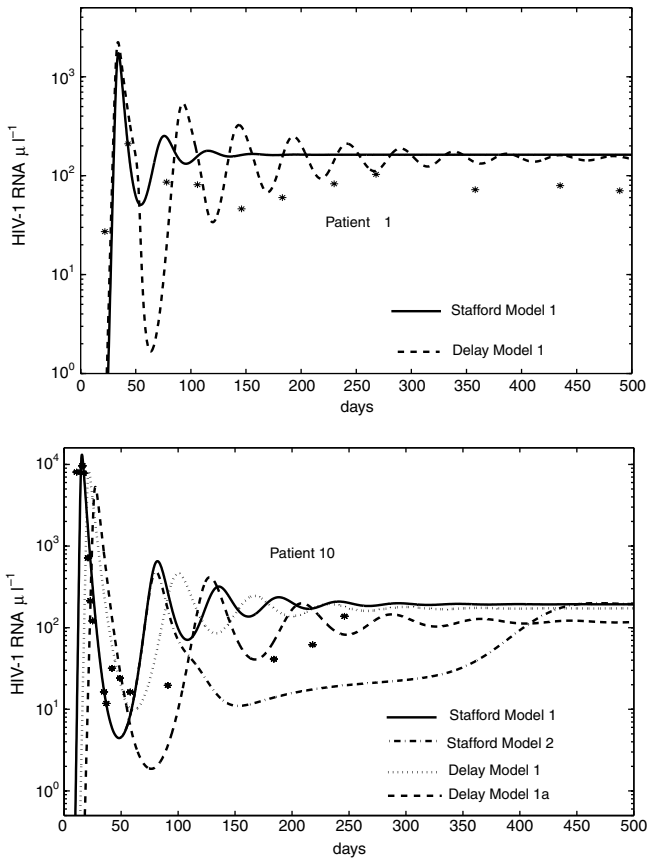

Fig. 5. For both patients 1 and 10 we find damped oscillations when using the estimated values from the Monte Carlo simulations suggesting a much different biological interpretation. Again, if we adjust the value of d_x in patient 10, from $d_x = 0.241$ viral-day^{-1} (Delay model 1) to $d_x = 2.241$ viral-day^{-1} (Delay model 1a) and change c from 5.19 day^{-1} to 7.19 day^{-1}, we find a fit that can mimic the model in Stafford that allows for an immune response (Stafford model 2). Asterisks ($*$) are patient data.

Carlo simulations and again found differences in the estimated parameters. The Delay model 2 (21) estimated δ to be $\delta = 0.134\,\text{day}^{-1}$ and c to be $1.559\,\text{day}^{-1}$ while the Delay model 1 (12) estimated $\delta = 0.514\,\text{day}^{-1}$ and $c = 0.557\,\text{day}^{-1}$. This is a reduction of 74% in δ and a 180% increase in c. However, as mentioned earlier, the data is fit equally well, suggesting the data does not statistically support one model over another. In most cases, we found that the Delay model 2 (21) dampened the oscillations more rapidly then in the non-logistic case (see Fig. 6).

We are trying to fit a model with many parameters, and it is apparent that more data points are needed to determine which model is more accurate with statistical significance. It is particularly striking that even with up to 16 data points available for testing model accuracy, these observations are consistent with more then one different model, each implying very different pathogenic behavior.

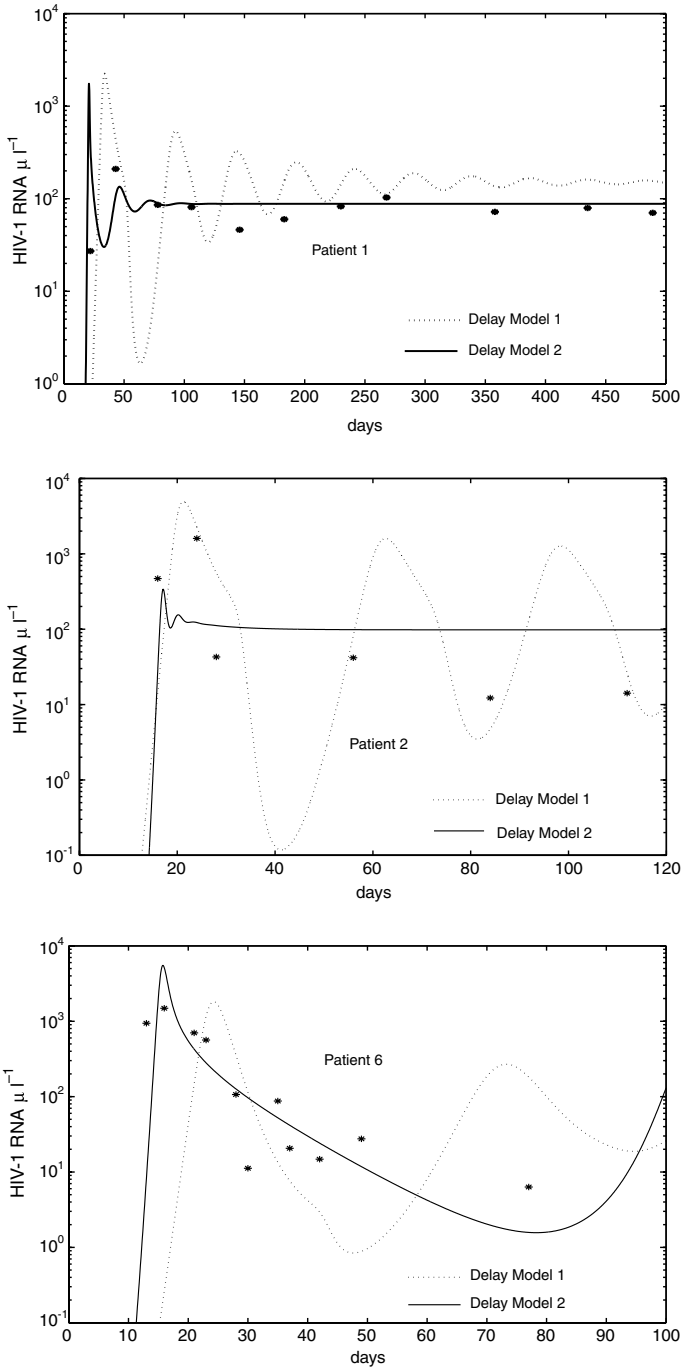

Fig. 6. Plots for patients 1, 2 and 6 comparing (12) (Delay model 1) with (21) (Delay model 2). For each patient, the corresponding parameters used are given in Tables 2 and 4. The logistic model seems to dampen the oscillations more rapidly, while still providing a reasonable fit. Asterisks (*) are patient data.

Acknowledgments

The authors would like to thank the anonymous reviewers for their insightful comments. The research of P.W.N. was supported in full by a Career Award at the Scientific Interface from the Burroughs Welcome Fund. M.S.C., D.M.B., and B.L.B. also acknowledge support from Burroughs Welcome and from a NSF Graduate Research Fellowship (B.L.B.).

References

[1] M. Stafford, L. Corey, Y. Cao, E. Daar, D. Ho, A. Perelson, Modeling plasma virus concentration during primary infection, J. Theor. Biol. 203 (2000) 285.
[2] R.J. De Boer, A.S. Perelson, Target cell limited and immune control models of HIV infection: a comparison, J. Theor. Biol. 194 (1998) 201.
[3] A. McLean, M. Nowak, Models of interactions between HIV and other pathogens, J. Theor. Biol. 155 (1992) 69.
[4] J. Murray, G. Kaufmann, A. Kelleher, D. Cooper, A model of primary HIV-1 infection, Math. Biosci. 154 (1998) 57.
[5] M. Nowak, A. Lloyd, G. Vasquez, T. Wiltrout, L. Wahl, N. Bischofberger, J. Williams, A. Kinter, A. Fauci, V. Hirsch, J. Lifson, Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection, J. Virol. 71 (1997) 7518.
[6] M. Nowak, R. May, R. Anderson, The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease, AIDS 4 (1990) 1095.
[7] A. Phillips, Reduction of HIV concentration during acute infection: independence from a specific immune response, Science 271 (1996) 497.
[8] M. Nowak, C. Bangham, Population dynamics of immune response to persistent viruses, Science 272 (1996) 74.
[9] T. Denny, B. Jensen, E. Gavin, A. Louzao, F. Vella, J. Oleske, W. Wong, Determination of CD4 and CD8 lymphocyte subsets by a new alternative fluorescence immunoassay, Clin. Diag. Lab. Immunol. 2 (1995) 330.
[10] M. Roederer, J. Dubs, M. Anderson, P. Raju, L. Herzenberg, CD8 naive T-cell counts decrease progressively in HIV-infected adults, J. Clin. Invest. 95 (1995) 2061.
[11] R. Mercado, S. Vijh, S. Allen, K. Kerksiek, I. Pilip, E. Pamer, Early programming of *T* cell populations responding to bacterial infection, J. Immunol. 165 (2000) 6833.
[12] P. Wong, E. Pamer, Antigen-independent CD8 *T* cell proliferation, J. Immunol. 166 (2001) 5864.
[13] M. van Stipdonk, E. Lemmens, S. Schoenberger, Naive CTLs require a single brief period of antigenic stimulation for clonal expansion and differentiation, Nat. Immunol. 2 (2001) 4239.
[14] R. Antia, C. Bergstrom, S. Pilyugin, S. Kaech, R. Ahmed, Models of CD8+ responses: What is the antigen-independent proliferation program, J. Theor. Biol. 221 (2003) 585.
[15] M. van Stipdonk, G. Hardenberg, M. Bijker, E. Lemmens, N. Droin, D. Green, S. S.P., Dynamic programming of CD8+ T lymphocyte responses, Nat. Immunol. 4 (2003) 361.
[16] Z. Dai, B. Konieczny, F. Lakkis, The dual role of IL-2 in the generation and maintenance of CD8+ memory *T* cells, J. Immunol. 165 (2000) 3031.
[17] M. Betts, D. Ambrozak, D. Douek, S. Bonhoeffer, J. Brenchley, J. Casazza, R. Koup, L. Picker, Analysis of total human immunodeficiency virus (HIV)-specific CD4+ and CD8+ T-cell responses: relationship to viral load in untreated HIV infection, J. Virol. 75 (2001) 11983.
[18] J. Forde, P. Nelson, Application of sturm sequences to bifurcation analysis of delay differential equations, J. Math. Anal. Appl. 300 (2004) 273.
[19] D. Bortz, P. Nelson, Sensitivity analysis of nonlinear lumped parameter models of HIV infection dynamics, Bull. Math. Biol. 66 (2004) 1009.
[20] H. Banks, B. Fitzpatrick, Statistical method for model comparison in parameter estimation problems for distributed delays, J. Math. Biol. 28 (1990) 501.

[21] P. Nelson, J. Mittler, A. Perelson, Effect of drug efficiency and the eclipse phase of the viral life cycle on estimates of HIV-1 viral dynamic parameters, J. AIDS 26 (2001) 405.
[22] A. Perelson, P. Nelson, Mathematical models of HIV dynamics in vivo, SIAM Rev. 41 (1999) 3.
[23] A. Perelson, A. Neumann, M. Markowitz, J. Leonard, D. Ho, HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time, Science 271 (1996) 1582.
[24] D. Bortz, P. Nelson, Model selection and mixed-effects modeling of HIV infection dynamics, Bull. Math. Biol., in press.