



A Stochastic Model for Early HIV-1 Population Dynamics

HENRY C. TUCKWELL*†‡ AND EMMANUELLE LE CORFEC*

* *Epidémiologie et Sciences de l'Information, Université Paris 6, INSERM U444, Institut Fédératif de Recherche sur la Santé St-Antoine, 27 rue Chaligny, F-75571 Paris Cedex 12, France* and † *Department of Mathematics, University of California, Irvine, CA 92697, U.S.A.*

(Received on 3 November 1997, Accepted in revised form on 28 July 1998)

A simple stochastic mathematical model is developed and investigated for early human immunodeficiency virus type-1 (HIV-1) population dynamics. The model, which is a multi-dimensional diffusion process, includes activated uninfected CD4⁺ T cells, latently and actively infected CD4⁺ T cells and free virions occurring in plasma. Stochastic effects are assumed to arise in the process of infection of CD4⁺ T cells and transitions may occur from uninfected to latently or actively infected cells by chance mechanisms. Using the best currently available parameter values, the intrinsic variability in response to a given initial infection is examined by solving the stochastic system numerically. We estimate the statistical distributions of the time of occurrence and the magnitude of the early peak in viral concentration. The maximum of the viral load has a value in the experimental range and its time of occurrence has a 95% confidence interval from 19.4 to 25.1 days. The stochastic nature of the growth of viral density is extremely pronounced in the first few days after initial infection. Threshold effects are noted at virion levels of about $3\text{--}5 \times 10^{-5} \text{ mm}^{-3}$. In addition to modeling the intrinsic variability in HIV-1 growth, we have explored the effects of perturbations in the parameter values in order to assess the additional stochastic effects of between-patient variability. We found that changes in the initial number of virions or dose size, the rate at which latently infected CD4⁺ T cells are converted to the actively infected form and the fraction of latent cells had only minor effects on the size, speed and variability of the response. In contrast, decreased speed and magnitude but greater variability in response were obtained when the death rate of uninfected CD4⁺ T cells, the death rate of actively infected cells and the clearance rate of the virus were increased or when the appearance rate of uninfected CD4⁺ T cells, the number of virions produced by infected cells, the infection rate of CD4⁺ T cells and the initial number of uninfected activated CD4⁺ T cells were decreased. We also determined the distribution of the time to reach a given virion density. From this distribution the probability of detection of the virus as a function of time can be estimated. The numerical results obtained are in the range of experimental values and are discussed in relation to recently proposed detection and testing procedures.

© 1998 Academic Press

1. Introduction

In the early stages of HIV-1 infection symptomatic primary infection coincides with high

concentrations of virions in both plasma and peripheral blood mononuclear cells (Niu *et al.*, 1993). The viral density then declines rapidly within several days, with a concomitant increase in cytotoxic T lymphocytes and the subsequent

‡ Author to whom correspondence should be addressed.

appearance of specific anti-HIV-1 antibody (Pantaleo *et al.*, 1993a). Mathematical models for HIV-1 population dynamics have been used in attempts to understand the basic mechanisms involved in the evolution of the infection at the microscopic level and to ascertain the effects of drug treatments, particularly reverse transcriptase inhibitors and protease inhibitors (Perelson, 1989; Perelson *et al.*, 1993; Essunger & Perelson, 1994; Phillips, 1996; Herz *et al.*, 1996; Nowak & Bangam, 1996; Perelson *et al.*, 1996; Bonhoeffer *et al.*, 1997; Kirschner & Webb, 1996; Perelson *et al.*, 1997; Wein *et al.*, 1997; McLean *et al.*, 1991). However, such analyses have been essentially deterministic and hence limited to only one possible solution whereas in reality the systems under consideration are intrinsically stochastic. Here we include these hitherto neglected but significant chance mechanisms. If these are introduced in a biologically meaningful fashion, one should be able to estimate their contributions to the variability in the early time course of the viral load, which has not been possible with deterministic models. Furthermore, one can obtain predictions of the probability that HIV-1 levels reach certain levels as a function of time since infection. Such levels can correspond to thresholds in various tests for the detection of HIV-1 in blood. Furthermore, it is expected that a stochastic approach can provide a more accurate quantitative basis for evaluating the efficacy of drug treatments in infected host populations. It is emphasised that in this work we concentrate only on the early period (to several weeks) after infection and not on the later progression to the acquired immune deficiency syndrome which may follow.

Deterministic models of HIV-1 dynamics have included at least two components; the uninfected $CD4^+$ T-cell lymphocyte count and the density of virus-producing cells, as in Bonhoeffer *et al.* (1997). However, in the construction of a model for early HIV-1 dynamics one may consider several components as basic. We will focus attention on the development of HIV-1 populations within the blood compartment. We will assume, as have many authors using deterministic models (for example, Essunger & Perelson, 1994; Phillips, 1996) that a $CD4^+$ T cell cannot be infected unless it first becomes activated,

although it is known that $CD4^+$ T cells may become abortively infected and either remain as such, or progress to the activated state (Chun *et al.*, 1997). Basic components might therefore include two classes of uninfected $CD4^+$ T cells, namely activated and unactivated cells (Fauci, 1996) although there is much uncertainty regarding the time course of changes from the unactivated to activated states and regarding the relative proportions of these two classes of cells (Phillips, 1996). There are also cytotoxic T-lymphocytes which directly attack virus producing $CD4^+$ T cells (Pantaleo & Fauci, 1996). Nevertheless, some authors have found it useful to consider simple models which include only viral density, uninfected $CD4^+$ T-cell numbers and those of one class of infected $CD4^+$ T-cell population. This gives rise to three-component models as in Herz *et al.* (1996), Nowak & Bangham (1996) and Kirschner & Webb (1996). However, we have found that a reasonably accurate picture is obtained for early HIV-1 dynamics using a stochastic four-component model which distinguishes between latently and actively infected $CD4^+$ T cells as in McLean *et al.* (1991) and Phillips (1996).

A more complicated model for HIV-1 dynamics might incorporate two connected compartments, namely plasma and lymphoid tissue (Pantaleo *et al.*, 1993b; Perelson *et al.*, 1996; Cohen *et al.*, 1997). The mechanisms of viral growth are quite different in the lymphoid compartment and it is probable that approaches to equilibrium conditions may occur with different time-scales there from those in plasma. Furthermore, despite the fact that a relatively small fraction of HIV-1 proliferation is thought to take place in plasma, Phillips (1996) assumed that the volume he considered is either one of blood or a smaller volume of lymphoid tissue which contains the same number of $CD4^+$ T cells. We will adopt a similar approach and defer the development of a more complex model until more details are known about the dynamics of HIV-1 growth in the lymphoid compartment. The resulting model, with its focus on plasma variables, has the advantage of simplicity while retaining the essential features of the ongoing dynamical processes.

2. Description of the Model

Let $T(t)$ be the number of uninfected, activated CD4⁺ T cells at time t , and $V(t)$ be the number of free virions, where these quantities refer to \bar{V} mm³ of plasma, which may be taken as whole blood. The distinction is made between those infected CD4⁺ T cells which do not immediately participate in viral reproduction, called latently infected cells and denoted by $L(t)$ (McLean *et al.*, 1991; Phillips, 1996) and actively infected cells which are involved in ongoing viral emission and which are denoted by $I(t)$. In Table 1 we list the principal variables and parameters of the model.

Now, the infection of CD4⁺ T-cell lymphocytes by HIV-1 is governed by the binding of virus particles to receptors and such an infection mechanism implies collisions or close contacts. These are clearly random events and in any small time interval the probability of these contingencies must in fact be quite small, given the dimensions and numbers of cells and virions in plasma. (We note that there is the possibility of multiple virion attachments to the same CD4⁺ T cell.) In the following, upper case letters will be used to signify random variables and random processes. Then in $(t, t + \delta t]$, let the change in uninfected cell numbers be δT . Then if δt is small enough there are, in a simplified approach, the two possibilities that δT is -1 , with probability $k_1 TV \delta t$, or that $\delta T = (\lambda - \mu TV) \delta t$, with probability $1 - k_1 TV \delta t$. Here k_1 is a constant representing the probability per unit time per

virion of a “successful” collision with, and attachment to, an uninfected CD4⁺ T cell. Similarly one can evaluate the random changes δL , δI and δV in the latently and actively infected cell and virion numbers, respectively.

Although the numbers of particles are whole numbers, one may make continuous approximations for the various components. In a stochastic model, this can be easily done by determining the first and second infinitesimal moments of the components to obtain a diffusion approximation (see for example Tuckwell, 1995). A brief explanation is given in the Appendix. In this approximation the components satisfy the following stochastic differential equations which are similar to those in previous deterministic models (e.g. Perelson *et al.*, 1993; Phillips, 1996) but with additional noise terms:

$$dT = (\lambda - \mu T - k_1 TV) dt - \sqrt{(k_1 TV)} dW \quad (1)$$

$$dL = (k_1 p TV - \mu L - \alpha L) dt + \sqrt{(k_1 p TV)} dW \quad (2)$$

$$dI = [k_1(1 - p)TV + \alpha L - aI]dt + \sqrt{(k_1(1 - p)TV)} dW \quad (3)$$

$$dV = (cI - \gamma V - k_2 TV) dt - \sqrt{(k_2 TV)} dW. \quad (4)$$

Here W is a standard Wiener process (i.e. mean 0, variance t at time t). The deterministic part of the differential equation for T contains a linear term in T , as in Phillips (1996), and may be viewed as a combination of the death rate term and a linear approximation to a logistic term (Perelson *et al.*, 1993) of the form

$$rT \left(1 - \frac{I + L + T}{T_{max}} \right),$$

where r is a constant. Such an approximation is expected to be reasonable because $I + L + T \ll T_{max}$ in the early period. We point out that previous articles in which this approximation was not made were often concerned principally with the later dynamics of HIV-1. The units of W in (1) are (CD4⁺ T cells)^{1/2} days^{1/2}. We have used the same symbol W throughout (1)–(4) even though the dimensions of this

TABLE 1

Variables and parameters	
t	Time in days since initial infection
T	Activated uninfected CD4 ⁺ T-cell concentration
L	Latently infected CD4 ⁺ T-cell concentration
I	Actively infected CD4 ⁺ T-cell concentration
V	HIV-1 virion density
λ	Appearance rate of uninfected CD4 ⁺ T cells
μ	Net death rate of uninfected CD4 ⁺ T cells
k_1	Infection rate per virion
k_2	Infection rate per uninfected CD4 ⁺ T cell
p	Proportion of infected cells which are latent
α	Activation rate of latently infected cells
a	Death rate of actively infected cells
c	Rate of virion emission by infected CD4 ⁺ T cells
γ	Death or clearance rate of virions
τ	Fraction of activated uninfected CD4 ⁺ T cells

quantity change as this has avoided the introduction of further constants with the same numerical values. However, the constants k_1 and k_2 are numerically equal with different dimensions.

Stochastic effects arise by virtue of the nature of the interactions between free virions and uninfected CD4⁺ T cells. However, we have introduced another source of variability by making the transitions from uninfected cells to either latently or actively infected cells stochastic. That is, infection results in a latently infected cell with probability p or in an actively infected cell with probability $1 - p$ when an interaction occurs. There is a deterministic conversion of latently infected cells to the actively infected state. We incorporate the additional chance mechanism in the solutions by simulation as follows. Let T_n, L_n, I_n, V_n approximate the corresponding continuous variables at time $t = n\delta t$ where $n = 0, 1, 2, \dots$. Then we put

$$L_{n+1} = L_n + (k_1 B_n T_n V_n - \mu L_n - \alpha L_n) \delta t + \sqrt{k_1 B_n T_n V_n \delta t} N_n \quad (5)$$

where $\{B_n\}$ is a sequence of independent variables which take the value 1 with probability p and the value 0 with probability $1 - p$; and $\{N_n\}$ are independent normals with mean 0 and variance 1. Similarly for I_n .

We are ignoring the time delay between infection and virus release. Note that compared with several corresponding deterministic models, there is an extra term, $-k_2 TV$, in the equation for dV/dt to account for the fact that whenever a CD4⁺ T cell is infected, not only does T decrease, but V decreases. As in many deterministic studies (Perelson, 1989; Perelson *et al.*, 1993) we include only one variable for the viral density so that different viral strains are not distinguished or are averaged over (cf. Abundo & Rossi, 1994). Furthermore, in order to simplify the model system, the details of the dynamics of CD4⁺ T cells themselves are not included in this work. It is noteworthy that the parameter k_1 , which determines the magnitudes of the random fluctuations, is already contained in the deterministic model. However, its empirical effects and statistical estimation will be affected by the presence of the noise terms.

3. Results

We have utilized the fact that stochastic systems such as the above can be solved by simulation using a strong Euler method (Tuckwell & Lansky, 1997). We assumed a total blood volume of $\bar{V} = 5 \times 10^6 \text{ mm}^3$ and employed values for the parameters which are currently the best available. Activated uninfected CD4⁺ T cells are assumed in the first instance to number 200 (1000τ) of the total of approximately 1000 mm^{-3} (Phillips, 1996) and the values of λ and a were taken from the same reference. The value of μ follows from the assumed initial equilibrium condition $T(0) = \lambda/\mu$. The initial number of virions was taken as 2 for whole blood. The standard parameter values, for a plasma volume of 1 mm^3 were as follows:

$$\begin{aligned} \lambda &= 0.272 \text{ CD4 day}^{-1} \text{ mm}^{-3}, \\ \mu &= 0.00136 \text{ day}^{-1} \text{ mm}^{-3}, \\ k_1 &= 0.00027 \text{ day}^{-1} (\text{virion mm}^{-3})^{-1}, \\ k_2 &= 0.00027 \text{ day}^{-1} (\text{CD4 mm}^{-3})^{-1}, \\ p &= 0.1, \\ \alpha &= 0.036 \text{ day}^{-1} \text{ mm}^{-3}, \\ a &= 0.33 \text{ day}^{-1} \text{ mm}^{-3}, \\ c &= 100 \text{ virion CD4}^{-1} \text{ day}^{-1}, \\ \gamma &= 2 \text{ day}^{-1}, \\ \tau &= 0.2. \end{aligned}$$

In the simulations we found it expeditious to insert reflecting barriers at very small uninfected cell and virion numbers on account of the very small numerical values of these quantities at the beginning of infection. This prevented the random fluctuation terms from taking these variables to unphysical negative values and has a negligible effect on the numerical solutions. In most of the simulations a time step of 0.01 days (14.4 min) was employed. This was found to be small enough so that at periodic intervals of 2 days, from 0 to 60 days, the distributions of viral load were not significantly different when the time step was further reduced.

In Fig. 1 we show sample paths of (T, L, I, V) (10 trials) for the above stochastic model. Standard initial values for the dynamical system, based on a volume of 1 mm^3 , were chosen to be $T(0) = 200$, $L(0) = 0$, $I(0) = 0$ and $V(0) = 4 \times 10^{-7}$. Considerable variability is apparent in the time course of the solutions, although, as will

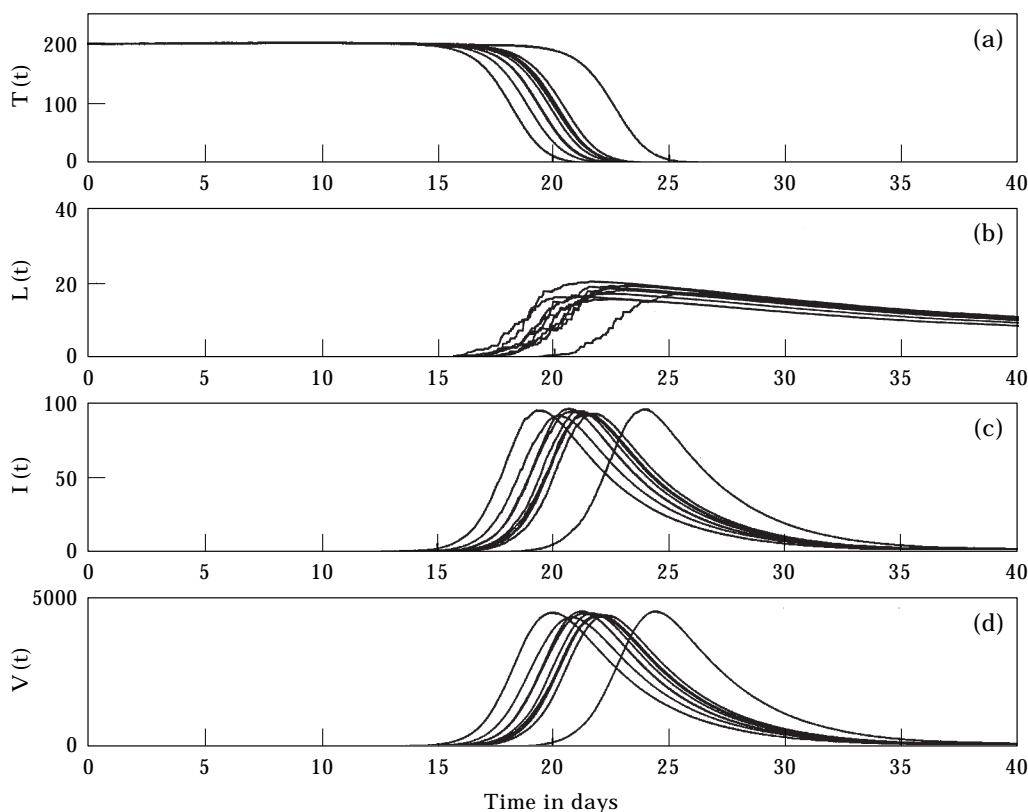


FIG. 1. Ten sample paths for the stochastic model of early HIV1 dynamics. Here are shown the time courses for densities of (a) T , uninfected $CD4^+$ T cells, (b) L , latently infected $CD4^+$ T cells, (c) I , actively infected, $CD4^+$ T cells and (d) V , plasma virions. Values of T , L , I , V are here given for 1 mm^3 . Parameter values are as in the standard set of Table 1 and initial values are $T(0) = 200$, $L(0) = 0$, $I(0) = 0$, $V(0) = 4 \times 10^{-7}$.

be elaborated upon below, once the viral density has reached appreciable levels, solutions are almost deterministic. The behavior of $V(t)$ here is similar to that observed experimentally (Piatak *et al.*, 1993; Baumberger *et al.*, 1993) and here shows considerable variability in regard to the maximal value, V_{max} , of $V(t)$. Although there are many paths which are close to the mean path, there are occasional outliers which coincide with a much slower time course for the development of the peak primary infection. The time courses of $I(t)$ and $V(t)$ are closely matched. The increasing phase of the paths of latently infected cell numbers, $L(t)$, is particularly stochastic, but the declining phase is very slow, smooth and almost linear.

3.1. DETAILS OF EARLY HIV-1 VIRAL DYNAMICS

Figure 2 shows the mean and standard deviation of the viral density as functions of time. The mean value of $V(t)$ only reaches a level

of about 3600 mm^{-3} which is far below the maxima of individual paths. The variability is greatest on the rising phase of $V(t)$ but declines as the maximum is approached. This can be seen from the values of the coefficient of variation (standard deviation divided by the mean) which are shown in the bottom part of Fig. 2. There is a subsequent mild increase in variability, followed by an approach to a regime of almost deterministic behavior as the intermediate term equilibrium is approached.

Figure 3 shows more details of the nature of the trajectories of the viral load. In the three parts of this figure the time and ordinate scales are different to emphasise different aspects of the temporal pattern of the evolution of viral load. The top part shows the same records as in Fig. 1. The middle part shows the early period between 1 and 3 days after initial infection by the virus. Here the paths are erratic and illustrate the stochastic nature of the infection of $CD4^+$ T cells

by virus particles. In the bottom part we show the time period from 8 to 10 days. By contrast, the trajectories are here very smooth and indicate that deterministic processes are dominating the growth process. Thus it seems that the development of primary infection divides into two regimes—a very early stochastic regime and a subsequent almost (quasi) deterministic one. A threshold-type effect is thus observed—note that in the middle part of Fig. 3, three trajectories are still at very low levels at $t = 3$ whereas the remaining seven paths have “taken off” and achieved values which are an order of magnitude greater. Thus when sufficiently high viral numbers are reached at some random time, a definite and practically certain large upsurge in virion level occurs. This is reminiscent of the phenomenon of threshold levels of voltage in action potential generation. If drug treatment is commenced before the time at which this threshold is reached, then virion numbers could

be contained at low levels with the possibility, in particular by random effects, that they become close to zero.

We also examined the time of occurrence, \bar{T}_{max} , and value V_{max} of the peak viral density. Figure 4 shows histograms for these quantities based on 200 simulations. The mean value of the maximal viral load (top part of Fig. 4) was 4494 mm^{-3} and its standard deviation was 104 mm^{-3} with a range of 4154 to 4709. These values are in the range of experimental values (Piatak *et al.*, 1993). The 95% range for V_{max} was (4263, 4677). The range of values of the time \bar{T}_{max} is from 19.1 to 27.0 days which is in good agreement with the observed times of occurrence of acute symptoms after infection (Niu *et al.*, 1993). The mean and standard deviation of \bar{T}_{max} were 21.5 and 1.4 days, respectively, with 95% range (19.4, 25.1). Note that for the corresponding deterministic model, there is only one solution, with a maximum virion density of

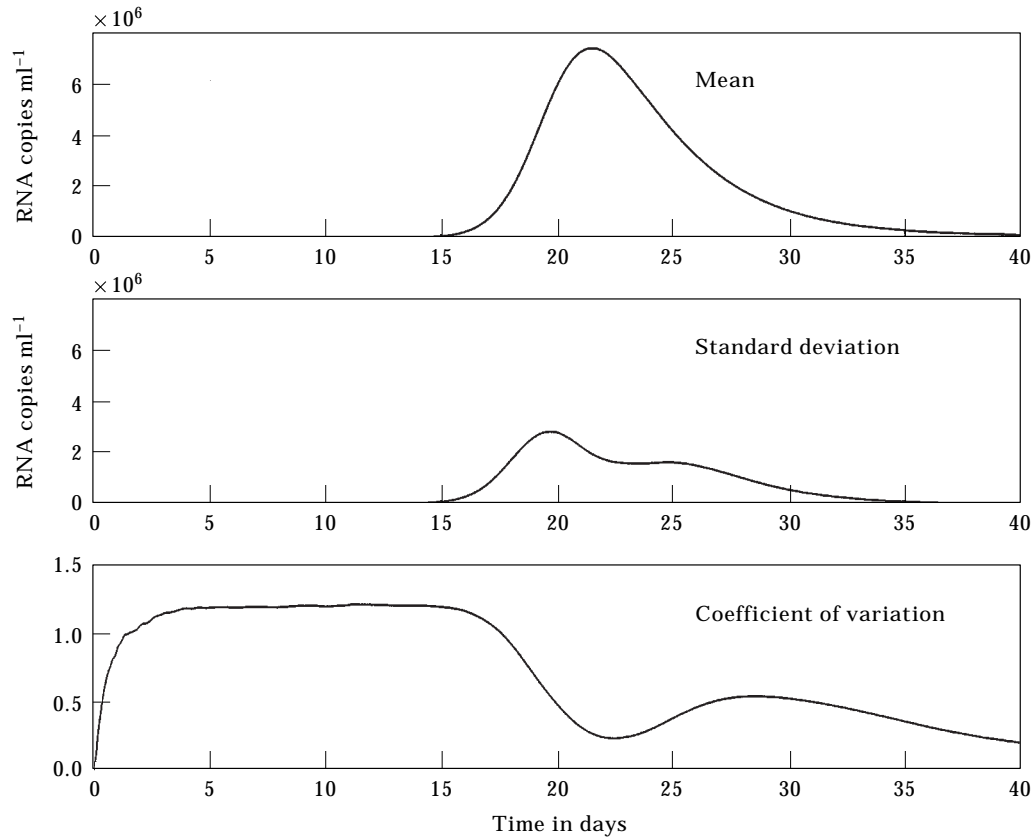


FIG. 2. The mean, standard deviation and coefficient of variation of the virion density as functions of time since initial infection up to 40 days. The parameters and initial values are the standard set. Viral load is here given as HIV1-RNA copies per ml.

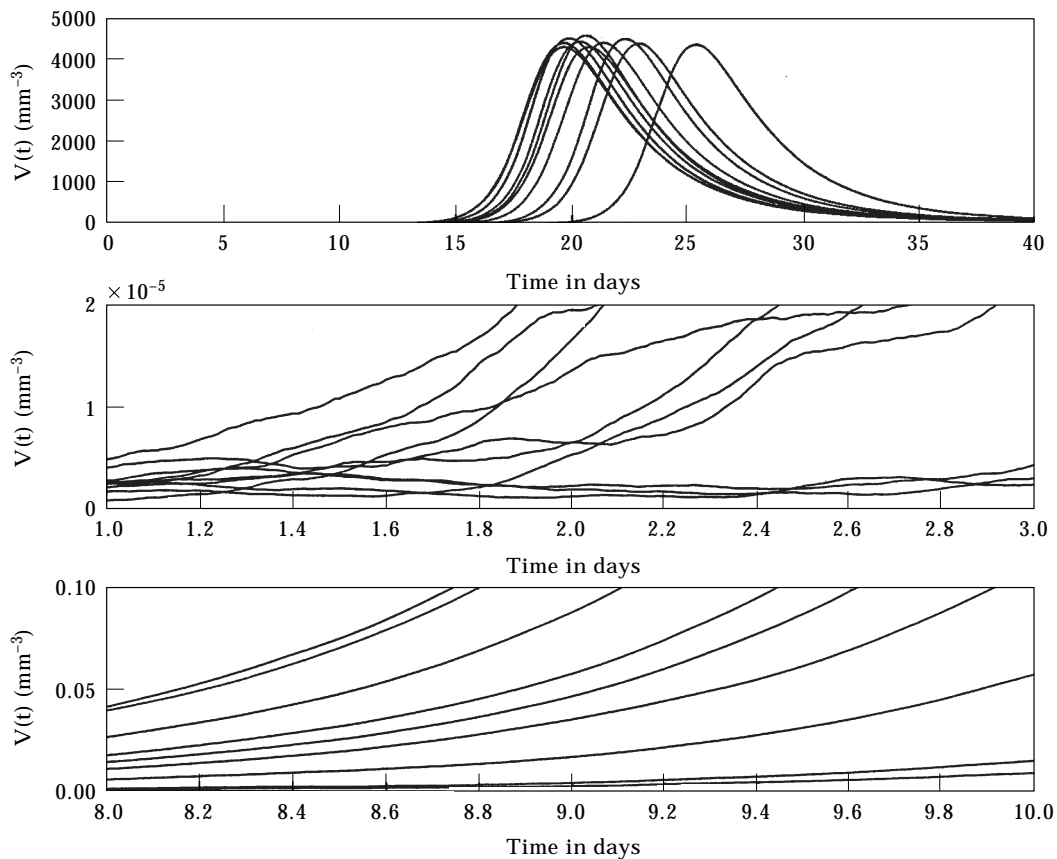


FIG. 3. Details of the evolution of the HIV-1 viral density $V(t)$. In the top part, results of 10 simulations for the number of virions per mm^3 are shown with standard parameter values. In the middle part, a magnified version of the early paths, from 1 to 3 days, is given, illustrating stochastic and threshold effects. The bottom part shows trajectories from 8 to 10 days, during which period the growth is practically deterministic.

4.571×10^3 occurring at 23.6 days after initial infection. For the stochastic model the variability in the time of occurrence and value of the peak viral load is not large as the coefficients of variation of these quantities are only 0.06 and 0.02, respectively.

3.2. EFFECTS OF CHANGES IN THE PARAMETERS

The values of the parameters and the initial values of the various components in the model are uncertain and also are expected to exhibit considerable variability amongst individuals. Hence we have systematically investigated the effects of changes in all these quantities. These effects are interesting in themselves but also provide a check that the magnitude of the random fluctuations is not a manifestation of the particular standard parameter set chosen. To this end, we have increased and decreased each parameter and certain initial values by 25%

relative to their standard values, employing a small number of simulations in each case.

Firstly we have found that these changes in λ , the appearance rate of activated uninfected CD4^+ T cells, $V(0)$, the initial number of virions, α , the rate at which latently infected CD4^+ T cells are converted to the actively infected form, and p , the fraction of infected cells which are latent, all had only minor effects on the solutions. These findings should not be interpreted as null effects for changes in these quantities in general, but only with respect to perturbations of the standard set. For example, if the parameter p is made much larger than 0.1, the time course of infection is very slow and the peak virion level much reduced. However, we did find that increasing $V(0)$ to as much as 10, or even 100, times the standard value had only a minor effect. This implies that in the contraction of the disease, the size of the initial viral

population is of little consequence in the subsequent growth of the viral population and thus in the severity or time course of symptoms. That is, infection by one or several particles leads to similar results.

The results of changes in the remaining parameters (those that had significant effects when their values are perturbed) are shown in Figs 5 and 6. In Fig. 5 the effects are shown of changing a , which is the rate of loss of actively infected $CD4^+$ T cells, and the parameter c which is proportional to the number of virions emitted by actively infected cells. It can be seen that a 25% increase in c and a 25% decrease in a led to greatly increased peak virion levels and more rapid growth. Changes in the opposite direction led to slower and more variable responses. In Fig. 6 are shown the simulation results for $V(t)$ as the remaining four parameters were varied. Increases in k_1 , the contact/infection rate of $CD4^+$ T cells by, for example virus particles, and τ , the fraction of uninfected $CD4^+$ T cells that are activated and thus able to be infected, led to

increased virion density at an earlier time with less variability. Note that increasing τ is equivalent to increasing $T(0)$ and that although the effects of a 25% change in k_1 were rather feeble, those which resulted from changing τ were somewhat dramatic, the total range for V_{max} being from about 3000 to 6000 mm^{-3} . The latter remark applies equally to decreased values in μ , the death rate of the uninfected cell population, and to γ , the clearance rate of free virions. It is observed that the remarks in the last two paragraphs in relation to the (mean) speed and (mean) amplitude but of course not the variability, would apply equally to the corresponding deterministic model.

We have also determined, from 200 simulations with the standard set of parameters, an estimate of the time, T_c , taken for the HIV-1 RNA level to reach certain specified values, V_c . In Fig. 7, the distributions of the times to reach the two assumed thresholds of detection, $V_c = 20$ (bottom histogram) and $V_c = 200$ (top histogram) copies ml^{-1} are shown. For the higher

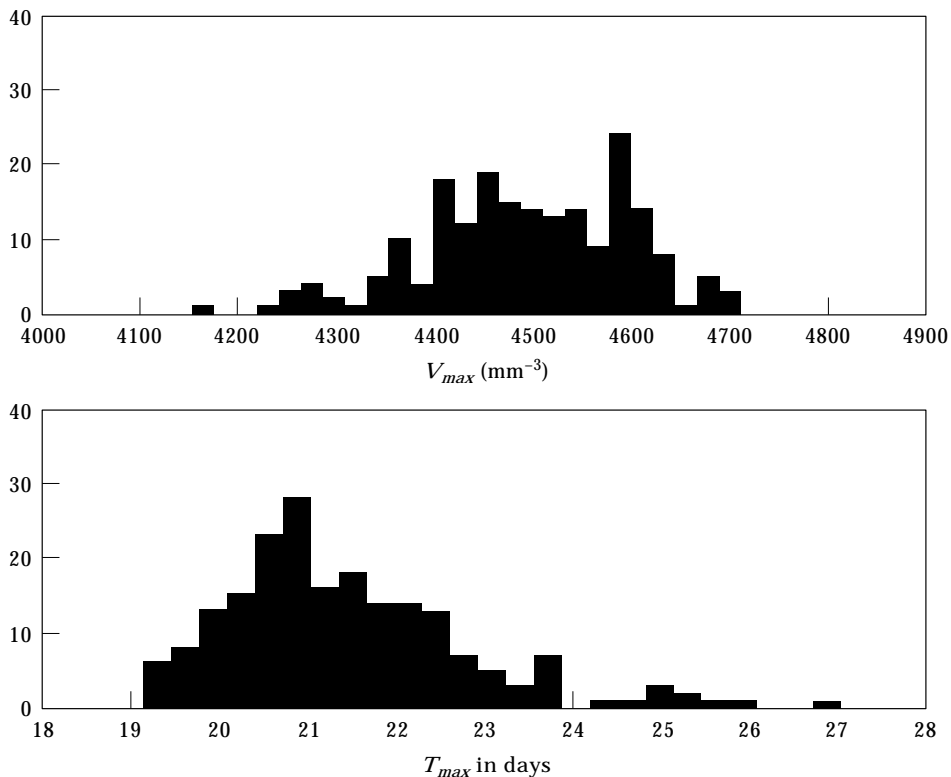


FIG. 4. The lower histogram shows the distribution from simulation of the time \bar{T}_{max} to reach the peak value V_{max} of V . The upper histogram shows the corresponding distribution for the magnitude of V_{max} . These results are based on 200 simulations with the standard parameter set.

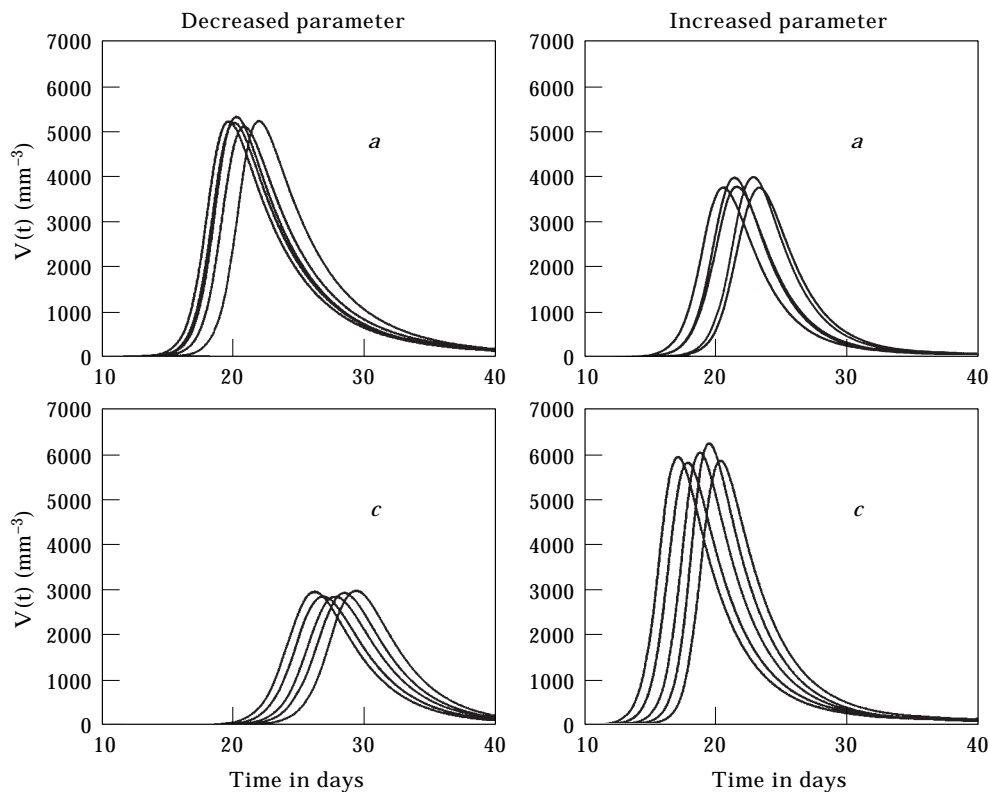


FIG. 5. Showing the effects of changes in the parameters of the model. Five solutions are shown for each new parameter set. Left figures show the results with a decreased value for a particular parameter whereas right figures show the results for an increase in the same parameter. All parameter changes were 25% relative to the standard value. Here are shown the paths of $V(t)$ obtained when a (top) and c (bottom) were varied.

threshold, the range of times to detection is from 8.3 to 16.3 days, with a 95% range from 8.6 to 14.2 days. The mean and standard deviation are 10.6 and 1.4 days, respectively. When the lower threshold is employed, the range of T_c is from 6.3 to 14.2 days, the 95% range of 6.6 to 12.2 days and the mean and standard deviation are 8.6 and 1.4 days, respectively. One may use this approach as a starting point in an attempt to ascertain the benefits of decreased thresholds in HIV-1 testing as for example in the testing of infected blood (Barbara, 1997; Legge, 1997).

4. Discussion and Conclusions

The above results point to the usefulness of the stochastic modeling of the growth of HIV-1 populations within infected individuals. We have considered the mechanisms which may contribute to the intrinsic variability during the infection process within a given host and also investigated systematically the variations which

might occur between patients in whom the parameters of the model may differ considerably. We have found that stochastic effects in the infection process have their origins predominantly in the early growth period, (from 0 to about 5 days after initial infection) when the virion density is relatively small. This observation has clear implications for the quantitative assessment of antiretroviral treatment in any period when the viral load has reached low levels. It is apparent that although a deterministic model may predict eradication of the virus, a stochastic treatment will more accurately predict a probability for such an eradication. For example, random fluctuations might take the viral density from a near-extinction level to higher levels.

Results for an accurate stochastic model are useful in the context of direct HIV-1 testing of donated blood (see Le Corfec *et al.*, 1998) and also for diagnoses of HIV-1 infection based on viral RNA measurements. Early diagnosis is

important so that antiretroviral treatment can be commenced as early as practicable (Ho, 1995). Stochastic modeling enables one to obtain a more accurate description of virion and CD4⁺ T-cell levels. Thus, as illustrated above, knowledge of the distribution of virion density as a function of time with a given test threshold can provide an estimate of the chance that a test will give a negative result, in the presence of HIV-1 infection, when the suspected time of primary infection is known or estimable. Furthermore, the recent demonstration (Finzi *et al.*, 1997; Wong *et al.*, 1997) that it is possible to isolate HIV-1 from peripheral blood mononuclear cells from patients on highly active antiretroviral therapy despite the suppression of their plasma HIV-1 RNA to very low levels for up to 2 years, indicates that stochastic modeling must be considered for studying viral dynamics under long term treatment.

Deterministic models also predict a uniquely specified time at which the viral load reaches any chosen level, whereas with a simple stochastic

model we can successfully predict the distribution of values for this quantity. Since the biological diagnosis of HIV-1 infection and consequent treatment can be based on measurements of plasma viral load (Busch *et al.*, 1995), questions may be addressed about the statistics of early detection such as those which relate to PCR (polymerase chain reaction) RNA testing.

We point out that in our simplified approach, we have omitted certain factors which may influence both the time course and the intrinsic variability of the growth of the HIV-1 population. One example, which was mentioned in the Introduction, is the inclusion of a lymphoid compartment. There is also the possibility of treating CD4⁺ T-cell dynamics by means of immigration and death processes and taking into account various mutant strains of the virus with birth and death processes (Abundo & Rossi, 1994). Another factor is that the number of virions emitted by an actively infected CD4⁺ T cell has been assumed to be a fixed number as has the half-life of infected cells, these quantities

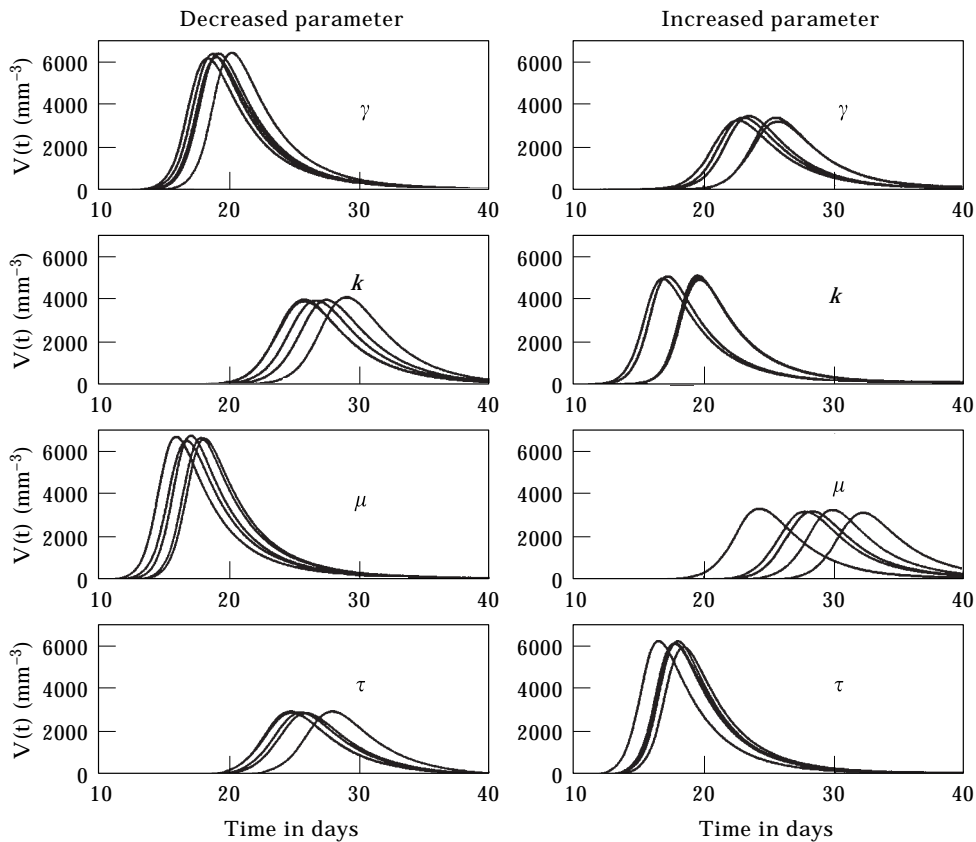


FIG. 6. As in Fig. 5, except here are shown (from top to bottom) the effects of varying the parameters γ , k , μ and τ .

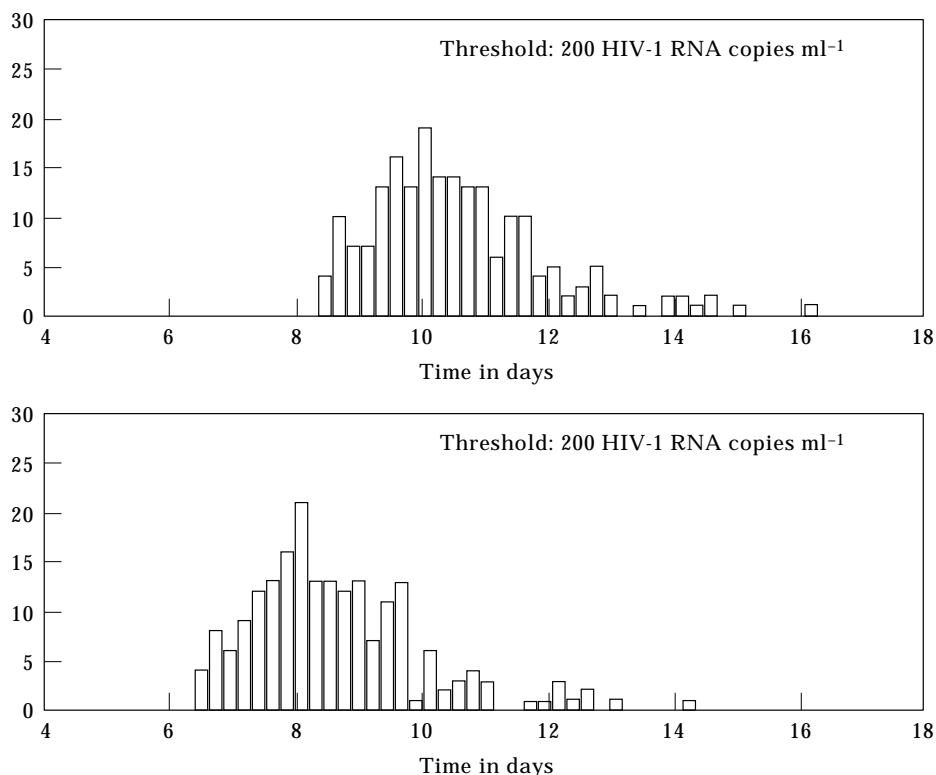


FIG. 7. Distributions based on 200 trials for the times (in days) to reach detection thresholds of 200 (top) and 20 (bottom) HIV-1 RNA copies ml^{-1} . These results were obtained with the standard parameter set.

being subsumed within the parameter c —in a fuller treatment such quantities could be made random. Similarly, the transition of latently infected cells to the actively infected state has been assumed to proceed at a constant rate, whereas such transitions are stochastic in nature. Moreover, even though the elimination of infected CD4^+ T cells by cytotoxic T cells, can be approximately subsumed within the parameter a , it is likely that a more complete treatment should include the details of the dynamics of the cytotoxic cells.

Finally, although we have focused on certain sources of intrinsic variability, between-patient sources may be easily incorporated by giving various parameters appropriate probability distributions. A suitable choice in initial studies is that each parameter is normally distributed with mean given by a best estimate value and a standard deviation of order indicated in the study of Perelson *et al.* (1997). One may perform simulations in which each parameter is sampled randomly from its (normal) distribution to obtain an estimate of the combined stochastic

effects of intrinsic mechanisms and between-patient variability. An indication of the magnitudes of these combined effects can be obtained from the results shown in Figs 5 and 6, though effects may not be additive. A more detailed study of such factors omitted in the present article will be undertaken in the near future.

We appreciate discussions with, and comments by, P.-Y. Boëlle, M. Burgard, F. Carrat, D. Costagliola, A. S. Perelson, C. Rouzioux and A.-J. Valleron. HCT thanks A.-J. Valleron and Institut National de la Santé et de la Recherche Médicale for a research professorship. ELC appreciates support from Agence Nationale de Recherche sur le SIDA and Fondation Pour la Recherche Médicale.

REFERENCES

- ABUNDO, M. & ROSSI, C. (1994). A stochastic model of the impact of genetic variability of HIV on the immune system in infected patients. In: *Modeling the AIDS Epidemic: Planning, Policy and Prediction* (Kaplan, E. H. & Brandeau, M. L., eds). New York: Raven.
- BARBARA, J. (1997). Benefit of using polymerase chain reaction to test blood donations will be considerable. *BMJ* **315**, 60.

- BAUMBERGER, C., KINLOCH-DE-LOËS, S., YERLY, S., HIRSCHL, B. & PERRIN, L. (1993). High levels of circulating RNA in patients with symptomatic HIV-1 infection. *AIDS* **7**, S59–S64.
- BONHOEFFER, S., COFFIN, J. M. & NOWAK, M. A. (1997). Human Immunodeficiency Virus drug therapy and virus load. *J. Virology* **71**, 3275–3278.
- BUSCH, M. P., LEE, L. L. & SATTEN, G. A. (1995). Time course of detection of viral and serologic markers preceding human immunodeficiency virus type 1 seroconversion: implications for screening of blood and tissue donors. *Transfusion* **35**, 91–97.
- CHUN, T.-W., CARRUTH, L., FINZI, D., SHEN, X., DIGIUSEPPE, J. A., TAYLOR, H., HERMANKOVA, M., CHADWICK, K., MARGOLICK, J., QUINN, T. C., KUO, Y.-H., BROOKMEYER, R., ZEIGER, M. A., BARDITCH-CROVO, P. & SILICIANO, R. F. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* **387**, 183–189.
- COHEN, O. J., PANTALEO, G., LAM, G. K. & FAUCI, A. S. (1997). Studies on lymphoid tissue from HIV-infected individuals: implications for the design of therapeutic strategies. *Springer Sem. Immunopathol.* **18**, 305.
- ESSUNGER, P. & PERELSON, A. S. (1994). Modeling HIV infection of CD4⁺ T-cell subpopulations. *J. theor. Biol.* **170**, 367–391.
- FAUCI, A. S. (1996). Host factors and the pathogenesis of HIV-induced disease. *Nature* **384**, 529–534.
- FINZI, D., HERMANKOVA, M., PIERSON, J., CARRUTH, L. M., BUCK, C., CLAUSON, R. E., QUINN, L. M., CHADWICK, K., MARGOLICK, J., BROOKMEYER, R., GALLANT, J., MARKOWITZ, M., HO, D. D., RICHMAN, D. D. & SILICIANO, R. F. (1997). Identification of a reservoir for HIV-1 in patients on a highly active antiretroviral therapy. *Science* **278**, 1295–1300.
- HERZ, A. V. M., BONHOEFFER, S., ANDERSON, R. M., MAY, R. M. & NOWAK, M. A. (1996). Viral dynamics *in vivo*: Limitations on estimates of intracellular delay and virus decay. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7247–7251.
- HO, D. D. (1995). Time to hit HIV, early and hard. *N. Engl. J. Med.* **333**, 450–451.
- KIRSCHNER, D. & WEBB, G. F. (1996). A model for treatment strategy in the chemotherapy of AIDS. *Bull. Math. Biol.* **58**, 367–390.
- LE CORFEC, E., LE PONT, F., TUCKWELL, H. C., ROUZIOUX, C. & COSTAGLIOLA, D. (1998). Direct HIV testing in blood donations: variation of benefit with detection threshold and pool size. (submitted).
- LEGG, A. (1997). Direct HIV testing of donated blood is inevitable. *BMJ* **314**, 1437.
- MCLEAN, A. R., EMERY, V. C., WEBSTER, A. & GRIFFITHS, P. D. (1991). Population dynamics of HIV within an individual after treatment with zidovudine. *AIDS* **5**, 485–489.
- NIU, M. J., STEIN, D. S. & SCHNITTMAN, S. M. (1993). Primary human immunodeficiency virus type 1 infection: review of pathogenesis and early treatment intervention in humans and animal retrovirus infections. *J. Infect. Dis.* **168**, 1490.
- NOWAK, M. A. & BANGHAM, C. R. M. (1996). Population dynamics of immune responses to persistent viruses. *Science* **272**, 74–79.
- PANTALEO, G. & FAUCI, A. S. (1996). Immunopathogenesis of HIV infection. *Ann. Rev. Microbiol.* **50**, 825–854.
- PANTALEO, G., GRAZIOSI, C. & FAUCI, A. S. (1993a). The immunopathogenesis of human immunodeficiency virus infection. *N. Engl. J. Med.* **328**, 327–335.
- PANTALEO, G., GRAZIOSI, C. & FAUCI, A. S. (1993b). The role of lymphoid organs in the immunopathogenesis of HIV infection. *AIDS* **7**, S19–S23.
- PERELSON, A. S. (1989). Modeling the interaction of the immune systems with HIV. In: *Mathematical and Statistical Approaches to AIDS Epidemiology* (Castillo-Chavez, C., ed.), Lectures in Biomathematics, Vol. 83, pp. 350–370. New York: Springer-Verlag.
- PERELSON, A. S., KIRSCHNER, D. E. & DE BOER, R. J. (1993). Dynamics of HIV infection of CD4⁺ T cells. *Math. Biosci.* **114**, 81–125.
- PERELSON, A. S., NEUMANN, A. U., MARKOWITZ, M., LEONARD, J. M. & HO, D. D. (1996). HIV-1 dynamics *in vivo*: viral 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. & 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: independence from a specific immune response. *Science* **271**, 497–499.
- PIATEK, M., JR., SAAG, M. S., YANG, L. C., CLARK, S. J., KAPPES, J. C., LUK-K-C., HAHN, B. H., SHAW, G. M. & LIFSON, J. D. (1993). High levels of HIV-1 plasma during all stages of infection determined by competitive PCR. *Science* **259**, 1749–1754.
- TUCKWELL, H. C. (1995). *Elementary Applications of Probability Theory: With an Introduction to Stochastic Differential Equations*. London: Chapman & Hall.
- TUCKWELL, H. C. & LANSKY, P. (1997). On the simulation of biological diffusion processes. *Comput. Biol. Med.* **27**, 1–7.
- WEIN, L. M., ZENIOS, S. A. & NOWAK, M. A. (1997). Dynamic multidrug therapies for HIV: a control theoretic approach. *J. theor. Biol.* **185**, 15–29.
- WONG, J. K., HEZAREH, M., GUNTARD, H. F., HAVLIR, D. V., IGNACIO, C. C., SPINA, C. A. & RICHMAN, D. D. (1997). Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* **278**, 1291–1295.

APPENDIX

The system of stochastic differential equations (1)–(4) can be conveniently rewritten, with the vector $\mathbf{X} = (X_1, X_2, X_3, X_4)'$ replacing $(T, L, I, V)'$, in the standard form

$$d\mathbf{X} = \alpha(\mathbf{X}) dt + \beta(\mathbf{X})d\mathbf{W},$$

where α is 4×1 , β is 4×4 , and $\mathbf{W} = (W_1, W_2, W_3, W_4)'$ is a four-vector of independent Wiener processes (primes denoting transpose).

The components of the vector α are the infinitesimal means

$$\alpha_k(y) \doteq \lim_{\delta t \rightarrow 0} \frac{E[\delta X_k | \mathbf{X}(t) = \mathbf{y}]}{\delta t}, \quad k = 1 - 4,$$

with $\mathbf{y} = (y_1, y_2, y_3, y_4)'$, $\delta X_k = X_k(t + \delta t) - X_k(t)$, and the components of the matrix β are the infinitesimal covariances or second moments

$$\beta_{kl}(y) \doteq \lim_{\delta t \rightarrow 0} \frac{E[\delta X_k \delta X_l | \mathbf{X}(t) = \mathbf{y}]}{\delta t}, \quad k, l = 1 - 4.$$

These quantities are easily evaluated from first principles as in Section 2. For example, the infinitesimal mean for the number of uninfected activated CD4⁺ T cells is

$$\begin{aligned} \alpha_1(y) &= \lim_{\delta t \rightarrow 0} \\ &\times \frac{[-k_1 y_1 y_4 \delta t + (1 - k_1 y_1 y_4 \delta t)(\lambda - \mu y_1) \delta t]}{\delta t} \\ &= \lambda - \mu y_1 - k_1 y_1 y_4. \end{aligned}$$

Instead of using simulation, one may proceed analytically to solve the model equations. Thus the four-component process, \mathbf{X} , also has an associated transition probability density function P which satisfies the forward Kolmogorov equation

$$\begin{aligned} \frac{\partial P}{\partial t} &= - \sum_{k=1}^4 \frac{\partial}{\partial y_k} (\alpha_k(y) P) + \frac{1}{2} \sum_{k=1}^4 \sum_{l=1}^4 \\ &\times \frac{\partial^2}{\partial y_k \partial y_l} ([\beta(y) \beta'(y)]_{kl} P). \end{aligned}$$

Evaluating the coefficients we find, with $A = y_1 y_4$ and $B = AP$,

$$\begin{aligned} \frac{\partial P}{\partial t} &= - \left[\frac{\partial}{\partial y_1} \{(\lambda - \mu y_1 - k_1 A) P\} \right. \\ &+ \frac{\partial}{\partial y_2} \{(k_1 p A - \mu y_2 - \alpha y_2) P\} \\ &+ \frac{\partial}{\partial y_3} \{(k_1(1-p)A + \alpha y_2 - \alpha y_3) P\} \\ &\left. + \frac{\partial}{\partial y_4} \{(c y_3 - \gamma y_4 - k_2 A) P\} \right] \\ &+ \frac{1}{2} \left[\frac{\partial^2}{\partial y_1^2} (k_1 B) + \frac{\partial^2}{\partial y_2^2} (k_1 p B) \right. \\ &+ \frac{\partial^2}{\partial y_3^2} (k_1(1-p)B) + \frac{\partial^2}{\partial y_4^2} (k_2 B) \left. \right] \\ &- \frac{\partial^2}{\partial y_1 y_2} (k_1 B) - \frac{\partial^2}{\partial y_1 y_3} (k_1 B) \\ &+ \frac{\partial^2}{\partial y_1 y_4} (\sqrt{k_1 k_2} B) \\ &+ \frac{\partial^2}{\partial y_2 y_3} (k_1 \sqrt{p(1-p)} B) \\ &- \frac{\partial^2}{\partial y_2 y_4} (\sqrt{k_1 k_2 p} B) \\ &- \frac{\partial^2}{\partial y_3 y_4} (\sqrt{k_1 k_2 (1-p)} B). \quad (\text{A.1}) \end{aligned}$$

This equation may be solved by numerical methods to give P , from which relevant statistical properties can be found. Furthermore, on setting $\partial P / \partial t = 0$ in (A.1), we obtain the equation satisfied by the steady-state probability density, $P^*(y)$, which gives the equilibrium distribution of \mathbf{X} at $t = \infty$.