Chapter 35

The Analysis of HIV Dynamics Using Mathematical Models

Ruy M. Ribeiro and Alan S. Perelson

The use of mathematical modeling and computer simulation in the study of infectious disease and, in particular HIV infection, has proven quite fruitful. The study of the dynamics of HIV infection represents a paradigm of success in the application of mathematical models to further our knowledge of disease pathogenesis (1,2). With these analyses, it was possible to quantify the rapidity of HIV infection and replication, the rate of virion clearance, the lifespan of productively infected cells (3–5), and predict the impact of treatment and the appearance of drug resistant mutations (6). Moreover, models have helped clarify controversial issues relating to the mechanism of T cell depletion in HIV infection and motivated new experimental and clinical studies.

In this chapter we will introduce the basic model for the analysis of HIV viral load decay under treatment, and explain the quantitative results obtained by these models when compared with experimental data. In later sections we will discuss extensions of this basic model to the analysis of viral reservoirs, primary infection and immune responses. Finally, we will consider models of cellular turnover, which are closely related to pathogenesis of HIV infection.

BASIC MODEL OF HIV INFECTION

Throughout most of the asymptomatic phase of HIV infection the plasma viral load remains approximately constant, at the set point reached after primary infection. This observation implies that the total amounts of virus produced and cleared in the body are balanced, i.e. production = clearance. If one of the processes in that equation dominated the other, than we would either have uncontrollable growth of virus (if production > clearance), or the virus would be eradicated (if production < clearance). This is a very powerful observation, because it means that by perturbing one side of the equation we can gain insight into the other. Thus, if we decrease the production rate, for example with antiretroviral therapy, then clearance dominates and the viral load decreases. This is exactly what is observed. Importantly, the magnitude and speed of the decrease allows the calculation of the clearance rate. Moreover, once the clearance rate is estimated, then the baseline production rate before therapy was initiated can also be estimated since at baseline, production = clearance. To obtain these estimates we need appropriate models, because the profile of the viral decay curve depends on several factors such as the clearance rate, the efficacy of the drug, and the death rate of virus producing cells. These models, based on the lifecycle of the virus, can then be used to fit quantitative data about viral decay under antiretroviral therapy.

We begin by considering a model of the viral lifecycle. Although HIV infects different types of cells, with variable replication rates, there is a consensus that the bulk of HIV production occurs in activated CD4+ T cells (7). Thus a basic model of HIV infection only includes activated cells ($T^+$), productively infected cells ($T^*$), and free virus ($V$). It is assumed that cells become activated at constant rate $\lambda$, and die at rate $d$ per cell. In the uninfected individual, for most of the time, there is a steady level of these cells, which is given by $\lambda/d$. However, in infected patients these cells are susceptible to infection by HIV at a rate proportional to the available numbers of uninfected cells and free virus, with rate constant $\kappa$. Thus cells become infected at rate $\kappa VT$, which corresponds to a mass-action term, common in chemical kinetics. Infected cells are created at this same rate, and die at rate $\delta$ per cell,
which may be different from the death rate of uninfected cells \((d)\), possibly higher. Finally, the virus is produced from the infected cells at rate \(p\), and it is cleared from the circulation at rate \(c\). Figure 35.1 represents this system in schematic form, and the corresponding equations are:

\[
\frac{dT}{dt} = \lambda - dT - kVT \tag{1}
\]

\[
\frac{dT^*}{dt} = kVT - \delta T^* \tag{2}
\]

\[
\frac{dV}{dt} = pT^* - cV \tag{3}
\]

We note that these terms and rates represent abstractions of possibly many different biological processes. Thus, for example, the clearance rate, \(c\), indicates how fast virus is cleared from the circulation, whatever the mechanisms—filtration by the reticuloendothelial system, opsonization, binding to follicular dendritic cells, phagocytosis, etc.

During the long asymptomatic phase of untreated HIV infection the viral load remains unchanged (8). This suggested initially that HIV replicated slowly. However, there was no quantitative way of measuring the viral turnover. Eqs. (1)–(3) show that the viral load can be at a steady state, i.e. unchanged in time \((dV/dt=0)\), as long as the production of virus \((pT^*)\) and the clearance \((cV)\) are balanced \((pT^* = cV)\). However, this balance may occur for many different sets of parameters, and thus viral production could be high, if clearance were also large, or those rates could both be small. Knowledge of the rate of production also indicates the burden imposed on the immune system and the opportunity for the virus to mutate.

**Treatment of HIV Infection**

When treatments with potent antiretroviral drugs, either protease inhibitors or a combination of a protease inhibitor and reverse transcriptase inhibitors, were administered to HIV-1 infected patients, a 1 to 2 log drop in the viral load was observed over the first one to two weeks of therapy. This observation allowed the quantification of the dynamics of virus production and clearance, using model (1)–(3) (3–5). Indeed, drug treatment represents a perturbation of the equilibrium by stopping the production of new virus, and the observed changes in viral load with time \((dV/dt)\) are an indirect measurement of the clearance rate of free virions.

To analyze such drug treatment experiments, a slight modification of model (1)–(3) is necessary. The drugs belonging to the class of reverse transcriptase inhibitors (RTI) reduce the ability of the virus to infect new cells, by interfering with the necessary process of reverse transcription of viral RNA into proviral DNA. Thus if the efficacy of these drugs is \(e_{RT}\), then only a proportion \(1 - e_{RT}\) of new infections will be productive and the infection terms in Eqs. (1)–(3), \(kVT\), have to be reduced to \((1 - e_{RT})kVT\). On the other hand, drugs belonging to the class of protease inhibitors (PI) interfere with the maturation of new virions, rendering them non-infectious. However, these non-infectious viral particles are still quantified by the commonly used RT-PCR based assays of viral load. Thus, it is important to follow the dynamics of these non-infections particles \((V_{NI})\). If the protease inhibitor efficacy is \(e_{PI}\), then only a proportion \(1 - e_{PI}\) of new viruses will be infectious, and a proportion, \(e_{PI}\), of viral particles will be non-infectious. The equations for the analysis of drug treatment data are then (Fig. 35.2):

\[
\frac{dT}{dt} = \lambda - dT - (1 - e_{RT})kVT \tag{4}
\]

\[
\frac{dT^*}{dt} = (1 - e_{RT})kVT - \delta T^* \tag{5}
\]

\[
\frac{dV_I}{dt} = (1 - e_{PI})pT^* - cV_I \tag{6}
\]

\[
\frac{dV_{NI}}{dt} = e_{PI}pT^* - cV_{NI} \tag{7}
\]
where $V_i$ represents infectious (or more precisely mature protease inhibitor untreated) virus, and $V_{NI}$ represents non-infectious (or immature) virus.

With Eqs. (4)–(7) it is in principle possible to estimate the decay rate of free virus ($c$) and infected cells ($\delta$), and from these the respective half-lives: $\ln(2)/c$ and $\ln(2)/\delta$. (Here ln(2) represents the natural logarithm of 2 and it is approximately 0.693.) It may also be possible to estimate the efficacies of the drugs, $\delta_{RT}$ and $\delta_{PP}$. However, estimates of all these parameters demand frequent sampling of viral loads, from the start of therapy, which is not always possible due to ethical considerations.

**Experimental Results**

The model in Eqs. (4)–(7) was used to fit experimental data of viral decay under drug treatment in HIV infection. With a 100% effective PI regimen, and assuming that the uninfected cells ($T$) remain constant during the short period of analysis, the number of viral particles ($V_i + V_{NI}$) per ml of blood decays as:

$$V(t) = V_0e^{-ct} + \frac{cV_0}{c-\delta}\left[\frac{cV_0}{c-\delta}(e^{-\delta t} - e^{-ct}) - \delta te^{-ct}\right]$$

(8) where $V_0$ is the pre-treatment (baseline) steady state viral load.

Estimates obtained by this method indicated that the half-life of free virus is less than 6 h, and that of infected cells is less than 1.6 days (3). The estimates are upper bounds because they are based on the assumption that drug therapy was 100% effective. In reality therapy is not so effective and residual viral production occurs. Thus, residual virus that is not accounted for in the model must also be cleared, and hence the true rates of virus and infected cell loss must be higher than is estimated by assuming a perfect drug. In Fig. 35.3 we show typical fits of the model solution, given by Eq. (8), to data for three patients. Since the virus is cleared rapidly, frequent sampling is needed at the initiation of treatment. Indeed for this experiment blood samples were initially obtained every 2 h, and still the estimated free virus half-life of 6 h is approximately at the limit of detection since it depends on so few initial data points, and is in fact just an upper bound.

The half-life of free virus has also been estimated using a different experimental method: plasma apheresis (9). In this method plasma with suspended viral particles is removed from the patient, and the same volume of fluid is re-infused. This technique again induces a perturbation of the equilibrium, since the clearance of HIV is no longer just $cV$, as in model (1)–(3), but larger due to the apheresis process. Frequent measurements of the decline in viral load with knowledge of the rate of removal of the virus by apheresis allow the estimate of $c$, and thus the half-life of free virus. This method yielded half-lives of about 1 h, i.e. even shorter than those estimated by application of drug therapy.

A half-life can be biologically interpreted as the time it takes for 50% of the free virus (infected cells) to be removed from circulation. Since before treatment, the viral load is in equilibrium, it follows that the viral (infected cell) production has to equal the removal rate. Thus a half-life of 1 h indicates that a typical individual with a set-point viral load of $10^5$ HIV RNA copies/ml (or equivalently $1.5 \times 10^8$ HIV RNA in the 15 liters of body water in a 70 kg person) has a total body daily production of between $10^{10}$ and $10^{11}$ viruses. The total body load of productively infected CD4 + T cells was estimated to be about $4 \times 10^7$ (10). If these cells have a half-life of 1.0 day, then $2 \times 10^7$ CD4 + T cells are destroyed and produced each day. (This half-life of one day is a recent estimate obtained with a very potent four-drug protocol (11).)

These large turnovers have implications for the generation of mutations in the viral genome. The estimated mutation rate of HIV is $3.4 \times 10^{-3}$ mutations per base per replication cycle (12), and the viral genome is about $10^7$ bases. Thus, it can be estimated that, on average, all one-point mutations will be produced many times each day, and that a proportion of all viable two-point mutations will also be produced. Depending on the selective disadvantage of these mutants, many will be present at low levels in the patient’s viral population, allowing for further accumulation of mutations (13,14). This explains the quasi-species nature of HIV and its rapid evolution in vivo. One consequence of this evolution is the rapid generation of viral variants resistant to all the available drugs when used alone, and this provides a rationale for combination therapy.

Potent combination antiretroviral therapy leads to a sustained viral decay over long periods of time until the viral load becomes undetectable by conventional assays. However, after the first week or two, the decay slows down and it no longer corresponds to the decline of short-lived productively infected cells, which have a half-life of about 1.0 day.

**MODELS FOR THE SECOND PHASE OF DECAY**

In the basic model (previous section) HIV is only produced from productively infected CD4+ T cells. However, to understand the second phase of decline observed in patients undergoing HAART, models with additional sources of HIV had to be developed (15). After productively infected CD4+ T cells decline due to treatment, these other sources would become dominant and their slower clearance leads to the observed slower decline in the levels of HIV RNA. Different possible putative sources were considered in the context of modeling: long-lived productively infected cells, latently infected cells, and tissue reservoirs, such as follicular...
dendritic cells (FDC) or drug sanctuaries (i.e. cell populations or tissues where drug penetration is sub-optimal).

Data fitting to HIV RNA decline cannot distinguish between those possible sources, since the fits for different models are equally good. For example, one can not distinguish between long-lived productively infected cells or latently infected cells being the source of second phase of decay. However, fitting HIV RNA and infected cell decay simultaneously indicates that the second phase of decay is more likely due to long-lived productively infected cells. These cells were estimated to have an average half-life of about 14 days (15), whereas latently infected cells without integrated provirus were estimated to have a half-life of about 8.5 days (16). With these estimates, we have calculated that productively infected cells contribute 93–99% of the virions produced at steady state, before therapy, whereas long-lived infected cells contribute only 1–7%, and latently infected cells less than 1% (17). HIV infects not only CD4+ T cells but also other cell types such as macrophages and dendritic cells, which are candidates for the long-lived productively infected cell pool. However, it is difficult to eliminate other contributions to the second phase of decline, and instead of a

FIG. 35.3. Fit of Eq. (8) to data of HIV-1 treatment, for three individuals (A, B, C), studied in (3). The black dots are the data-points and the solid lines the theoretical fits, from Eq. (8).
long-lived infected cell pool the second phase of decay could actually correspond to release of virions trapped on follicular dendritic cells (FDC) (17) (Fig. 35.4). There are large amounts of potentially infectious virions attached to FDC, and these may be released in a biphasic mode, if binding is reversible and multivalent (18). Thus, it is conceivable that these virions are an important source of HIV RNA during drug treatment (19). If this is the case, then the estimates of viral clearance and infected cell death rates would have to be revised upward, because these rates would have to compensate not only for the residual production of virus from infected cells, but also for the release of virus from FDC. Moreover, virus may remain bound to FDC for prolonged periods of time, perhaps decades (18). If they maintain their infectiousness, then they represent a serious obstacle to current therapies, since they could re-seed a productive infection as soon as therapy is stopped.

Finally, detailed measurements of HIV provirus integrated in resting memory CD4+ T cells indicate that this cell population can constitute an inducible long-lived latent reservoir (18–21). Measurements of the decay of PBMC infectiousness by limiting dilution analysis suggests that this reservoir could lead to a third, even slower, phase of decay of plasma HIV RNA. Analysis of the decay of the latent reservoir in patients treated with potent combination therapy, with models that assume exponential decay of the reservoir, show that the half-life of cells carrying replication competent viral genomes varies between six and 44 months (18–21). Since, it is possible that this latent pool is replenished whenever there is active production of virus, for example due to temporary incomplete suppression of replication by the drug regimen, it is unclear whether the long half-lives in this range represent the true rate of decay of the reservoir or are artificially raised due to reseeding of the reservoir.

EXTENSIONS OF THE BASIC MODEL

Most of the models used to analyze HIV RNA decay do not take into account the immune response mounted against the virus. However, several lines of evidence point to the importance of cytotoxic T lymphocyte (CTL) control of the virus. These include viral evolution and escape under CTL pressure (22), the rise of viremia in CD8+ T cell depletion studies in macaques (23,24), and the concomitant appearance of a HIV-specific CTL response and the decline of virus in primary infection (25).

It is worth pointing out that models without an immune response can account for the general dynamics of virus during primary infection (26,27). In these models, the viral load increase and subsequent decrease is controlled by the availability of target CD4+ T cells for the virus to infect and spread (26). However, when data are analyzed over periods longer than 100 days after infection, in some patients the predicted viral set point by the basic model is larger than observed, suggesting an effect of immune control. A way to model immune control by CTL-mediated cytolysis is simply to assume that the death rate of infected cells (δ) in Eq. (2) is not constant (27). For example, Stafford et al. (27) modeled δ as the sum of a constant part, δ₀, and a part, δ₁(V), that increases with viral (antigen) load, i.e. δ = δ₀ + δ₁(V). The use of models that allowed an immune response provided a better fit to the viral load data obtained from some but not all primary infection patients studied by Stafford et al. (27). Changing the rate of death of productive infected cells is only one of a number of possible implementations of the effect of the immune response. Another possibility is that the antiviral effect of CD8+ T cells is due to secretion of soluble factors that inhibit viral production (28). In this case, it is the production of virus by infected cells, p, in Eq. (3) that
is not constant, but rather decreases as CD8+ T cells increase (1,27).

Models of immune response in HIV infection have also been developed in other contexts. For example, to explain the two phase initial decay of viral load under therapy (29), to explore the cytopathicity of HIV (30), to explain HIV pathology and pathogenesis (31), to study viral evolution and diversification (32), and the role of CD4+ T cell help in maintaining the immune response (33).

MODELS OF T-LYMPHOCYTE DYNAMICS

An important question in HIV infection has been what is the effect of infection in T-lymphocyte turnover. Several experimental methods have been used to try to look at this question (34–41). For the quantitative interpretation of these experimental results, the use of mathematical models became crucial. Here we will only refer to two methods, which attempt to directly quantify turnover of T cells. Both involve the incorporation of a label into the DNA of replicating T cells. In one the label used is bromodeoxyuridine (BrdU), and in the other it is deuterated glucose (D-glucose).

The most informative way to use these labels is in pulse-chase experiments with frequent measurements to assess the kinetics of acquisition and loss of label. For modeling purposes, these methods differ mainly in the way measurements are taken. In the BrdU case, the fraction of labeled cells is measured, whereas in the deuterated glucose case, the fraction of labeled DNA is measured (42). This subtle difference has important implications in the definition of the model. For example, the pool of unlabeled cells is reduced both by proliferation (cell acquires label) and death (cell is removed), when one measures cell labeling with BrdU (43). However, if one uses DNA labeling, then death of a cell will reduce the pool of unlabeled DNA, but proliferation of a cell does not reduce the pool of unlabeled DNA. This is because replication of DNA is accomplished by copying the original unlabeled DNA strand template, which thus is still present in the progeny (46). Thus, when an unlabeled DNA strand, U, is copied, the result is one unlabeled and one labeled, U + L, strands (Fig. 35.5). This difference between the two methods is easily seen in the equations used to analyze each of the experiments (see Table 35.1).

In Table 35.1, in the BrdU method, U and L refer to unlabeled and labeled cells; and in the D-glucose method, the same letters refer to unlabeled and labeled DNA, respectively. In these equations, d and p are the death and proliferation rates of lymphocytes, whereas sL and sU represent potential sources of unlabeled and labeled material (cells or DNA) into the proliferating pool, where measurements are taken. As can be seen in Table 35.1, the equations to describe D-glucose labeling can be reduced to expressions involving only the death rate d and the sources, whereas in the BrdU expressions both death and proliferation rates are present, in addition to the sources. Thus the D-glucose data can be fitted with a model with fewer parameters. Both models assume that the cell population being labeled is at steady state (T0), so that the total number of cells and total amount of DNA is constant. Thus, for each method U + L = constant and the fraction labeled, fL(t), can be determined from either U(t) or L(t), as shown in Table 35.2. In this table, T0 represents the total number of cells in the BrdU method, and the total amount of DNA in the case of D-glucose labeling.

The use of these types of models has shown that in untreated HIV infection the average proliferation and death rates are increased threefold or more, and that treatment reduces these rates (38–41). Moreover, this reduction is dependent on the time since the initiation of treatment and after one year of antiretroviral therapy, the values for the proliferation and death rates are nearly equal to those of uninfected subjects (39).

One drawback of these models is that to obtain good agreement with the data one needs to assume a source of unlabeled cells (DNA) into the labeling population. However, the exact nature of this source is not known, and it is indeed controversial (38,39,45). We have built a more realistic model, based on two populations of resting and dividing T cells, to study this issue (44,45). In this model, resting T cells can be activated into the dividing pool, and the death and proliferation rates in the activated population are the same. In this way, the model allows for the differentiation between the dynamics of proliferation and

### TABLE 35.1. Equations for the analysis of T cell labeling experiments

<table>
<thead>
<tr>
<th></th>
<th>BrdU method</th>
<th>D-glucose method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Labeling</strong></td>
<td>( \frac{dU}{dt} = s_U - pU - dU )</td>
<td>( \frac{dU}{dt} = s_U - dU )</td>
</tr>
<tr>
<td><strong>De-labeling</strong></td>
<td>( \frac{dL}{dt} = s_U + pL - dL )</td>
<td>( \frac{dL}{dt} = s_L - dL )</td>
</tr>
</tbody>
</table>

FIG. 35.5. Diagram of the labeling of T cells under two protocols: BrdU labeling of dividing cells, and D-glucose labeling of the DNA of dividing cells. Labeled cells or both strands of DNA labeled are shown as filled circles, and half-filled circles indicate that only half of the DNA strands are labeled. Notice that in the D-glucose case the amount of unlabeled/labeled DNA remains constant with division during the labeling/de-labeling period, respectively.
death, and the dynamics of activation. Surprisingly, although the data on CD4+ and CD8+ T cells, look similar application of the same model leads to very different conclusions. The analysis shows that in the CD8+ T cells the fraction of activated cells is increased in relation to healthy individuals, even though the proliferation and death rates of activated CD8+ T cells are the same in infected and uninfected people. On the other hand, in the CD4+ T cell compartment the proliferation and death rates of activated cells are increased in relation to healthy individuals. Even though these short-term experiments that assume equilibrium cannot explain the long-term dynamics of T cells in HIV infection, this result helps explain why CD4+ and CD8+ T cells show different behavior during the course of infection.

CONCLUSIONS

Modeling has become an important tool in the analysis and interpretation of experiments in HIV research. The use of these models has permitted the estimation of important kinetic parameters that allow us to understand the dynamics of HIV infection, to test hypotheses, and to make new predictions. Overall, the use of models has resulted in a significantly better understanding of viral dynamics, the appearance of drug resistant mutations, and the pathogenesis of infection (47). In the future, collaborations between experimentalists and modelers should lead to new ways to plan, conduct, and analyze experiments.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy and supported by N.I.H. grants RR06555, AI28433, and AI40387. Most of the experiments reported here were done in collaboration with David D. Ho, Martin Markowitz, and Hiroshi Mohri, Aaron Diamond AIDS Research Center, Rockefeller University, NY. Without these experiments and the intellectual collaboration of David Ho this work would never have been done.

REFERENCES


