# The Differential Infectivity and Staged Progression Models for the Transmission of HIV *†

James M. Hyman
Theoretical Division, MS-B284
Center for Nonlinear Studies
Los Alamos National Laboratory
Los Alamos, NM 87545

Jia Li
Department of Mathematical Sciences
University of Alabama in Huntsville
Huntsville, AL 35899

E. Ann Stanley
Theoretical Division, MS-B284
Los Alamos National Laboratory
Los Alamos, NM 87545

November 5, 1998

**Abstract**

Recent studies of HIV RNA in infected individuals show that viral levels vary widely between individuals and within the same individual over time. Individuals with higher viral loads during the chronic phase tend to develop AIDS more rapidly. If RNA levels are correlated with infectiousness, these variations explain puzzling results from HIV transmission studies and suggest that a small subset of infected people may be responsible for a disproportionate number of infections. We use two simple models to study the impact of variations in infectiousness. In the first model, we account for different levels of virus between

*This research was supported by the Department of Energy under contracts W-7405-ENG-36 and the Applied Mathematical Sciences Program KC-07-01-01.

†The authors thank John Jacquez and an anonymous referee for their valuable comments.

individuals during the chronic phase of infection, and the increase in the average time from infection to AIDS that goes along with a decreased viral load. The second model follows the more standard hypothesis that infected individuals progress through a series of infection stages, with the infectiousness of a person depending upon his current disease stage. We derive and compare threshold conditions for the two models and find explicit formulas of their endemic equilibria. We show that formulas for both models can be put into a standard form, which allows for a clear interpretation. We define the relative impact of each group as the fraction of infections being caused by that group. We use these formulas and numerical simulations to examine the relative importance of different stages of infection and different chronic levels of virus to the spreading of the disease. The acute stage and the most infectious group both appear to have a disproportionate effect, especially on the early epidemic. Contact tracing to identify superspreaders and alertness to the symptoms of acute HIV infection may both be needed to contain this epidemic.

**Keywords:** *Differential infectivity, staged progression, reproductive number, endemic equilibrium, stability, sensitivity*

# 1 Introduction

Newly developed techniques for measuring HIV RNA levels are allowing researchers to develop a picture of HIV infection patterns. HIV-1 RNA levels in plasma and serum become extremely high during the 1-2 weeks of acute primary infection, before there is a detectable immune response [1, 2]. These levels are higher than at any other time during infection. Acute primary infection is followed by a chronic phase. During the chronic phase, HIV RNA levels drop several orders of magnitude and remain "nearly constant" for years [3, 4, 5], where "nearly constant" includes fluctuations that are less than an order of magnitude up and down for about 90% of the cohort and less than a factor of 100 for the remaining [4]. Fluctuations may be caused by transient illnesses and vaccinations. Successful therapy causes a drop in the viral load to a new level that is maintained until viral resistance develops [6]. Viral levels differ by many orders of magnitude between individuals. Those people with high viral loads in the chronic phase tend to progress rapidly to AIDS, whereas those with very low loads tend to be slow or nonprogressors [4, 5, 7, 8]. During late chronic infection, there is a small increase in HIV-1 RNA levels, at most tenfold, in many individuals [3].

Common sense says that viral levels in serum and plasma are correlated with infectiousness. If this is the case, then these results on HIV-1 RNA levels can explain much of the data on HIV transmission in couples. Couples studies have found that some individuals transfer the infection to their sexual partners after only a few contacts, but other couples have had thousands of unprotected contacts without transferring infection [9, 10, 11, 12]. A few epidemiological studies for small cohorts have found that either a partner transferred the virus early in the course of infection, or it was not transferred at all [13]. Some researchers have found evidence for increased transmission late in infection [14, 15] although others have not [11, 13]. Sometimes late-stage transmission does not occur because of the increased use of protective methods among couples; however, late-stage transmission occurred infrequently in one study even when the use of protective methods was controlled for in the data analysis [11].

These couples studies show that there must be either great variability in the infectiousness among infected individuals or great variability in the susceptibility of their partners, or both. The HIV-1 RNA data support the idea of variations in infectiousness and suggest that there are differences of many orders of magnitude in viral shedding rates both over time and between individuals. In this paper, we focus on those possibilities, and neglect variations in susceptibility. This not only allows us to focus in on the impact of variations in transmissibility, and keeps the mathematics more tractable, but it is also easy to see that variations in susceptibility will not affect the dynamics of an epidemic until depletion of the most susceptible groups occurs. However, note that the CCR5 results [16, 17] indicate that some individuals are not susceptible to infection: since they are a small fraction of the actual population, few contacts will be with them and these individuals can be accounted for with our models by simply assuming they don't belong to the susceptible population.

Variations in infectiousness over time can be explained as part of a Markov chain, or staged-progression (SP), model in which infected individuals sequentially pass through a series of stages, being highly infectious in the first few weeks after their own infection, then having low infectivity for many years, and finally becoming gradually more infectious as their immune system breaks down and they progress to AIDS. This

Markov chain model also provides an explanation for the very low progression rates to AIDS in the first few years after infection and allows for a good fit to the data for the distribution of the time from infection to AIDS [18]. Many modelers and statisticians have studied the SP hypothesis (see [18, 19, 20, 21, 22, 23, 24, 25]), first proposed because of early studies indicating that viral load in the bloodstream increases late in infection, as individuals begin to show signs of impaired immunity [26, 27] and indications of virus in the bloodstream before there is an antibody response [28, 29].

In this paper we study this SP hypothesis further, using a simple model. However, the HIV-1 RNA data show that the SP hypothesis is incomplete. Infected individuals have different levels of virus after the acute phase, and those with high levels progress to AIDS more rapidly than those with low levels. We separate the issues by proposing a new model that only accounts for differences between infected people, and we refer it to as a differential infectivity (DI) model. In our simple DI model, individuals enter a specific group when they become infected and stay in that group until they are no longer involved in transmission. Their infectivity and progression rates to AIDS depend upon which group they are in. In a future paper, we will examine what happens when the two processes are combined into a DISP model in which individuals both go through stages and have intrinsic differences in viral loads and progression rates.

We derive explicit formulas for the reproductive numbers and the endemic steady states for the DI and SP models, including the fraction of infections being caused by each group at equilibrium. By properly defining the mean duration of infection and the mean transmissibility of infected individuals, we express all of our formulas for the reproductive number and endemic states in the same and easily interpreted form for both models. We determine a baseline set of parameters for both models and use these formulas and numerical simulations of the transient dynamics to study which groups in each model are causing the bulk of the infections at different points in time. For the SP model, this provides further insight into the results in [21, 30]. Their numerical simulations showed that when partner acquisition rates are high, the bulk of the infections early in the epidemic are caused by those in the acute infectious stage. Our results indicate that this is also the case at fairly moderate partner acquisition rates and that as the epidemic progresses, the late-stage becomes more important to disease transmission than the early acute stage. We also show that a small number of individuals who are highly infectious during the chronic stage can have a disproportionate impact on the epidemic, even if they have a short life expectancy.

Note that differential infectiousness was studied in [31] for diseases transmitted by casual contact, which have a different mathematical structure than sexually transmitted disease models. That study also considered the impact of superspreaders, but from a different perspective.

# 2 The DI Model

## 2.1 The Model Formulation

In order to examine only one question at a time, we assume that the susceptible population is homogeneous and we neglect variations in susceptibility, risk behavior, and many other factors associated with the dynamics of HIV spread. We also assume that the population we are studying is a small, high-risk subset of a larger population. The larger embedding population is relatively free of HIV and provides a constant source of uninfected individuals entering the high-risk population we are studying. For example, we might apply this model to the homosexual population of a major American city or to a group of highly active heterosexuals. When no virus is present in the population, the population of susceptible individuals, $S$, has a constant steady state, $S^0$. This equilibrium is thus assumed to be maintained by a constant inflow and outflow during which time each individual remains in the population an average of $\mu^{-1}$ years; where $\mu$ is the removal rate due to natural death in the absence of HIV infection, migration, and changes in sexual behavior. Individuals are infected by HIV at a per capita rate $\lambda(t)$.

The infected population is subdivided into $n$ subgroups, $I_1, I_2, \cdots, I_n$. Upon infection, an individual enters subgroup $i$ with probability $p_i$ and stays in this group until becoming inactive in transmission, where $\sum_{i=1}^{n} p_i = 1$. We assume that the infection subgroup is not a transmissible property of the HIV virus because there is no evidence to the contrary. In fact, one study has shown that some characteristics of the virus (resistance to AZT) can change between infector and infectee [1]. By treating the susceptible population as a homogeneous group, we also neglect, for simplicity, any significant links between susceptibility and infectiousness or progression rates that may occur due to human genetics such as CCR5.

The rate, $\nu_i$, of leaving the high-risk population because of behavior changes that are induced by either HIV-related illnesses or a positive HIV test, and subsequently the desire not to transmit infection, may depend on the subgroup, since there may be a link between the amount of virus being shed by an infected individual an how quickly an individual gets sick. Let $A$ denote this subgroup of removed people. People in $A$ are assumed to die at a rate $\delta \geq \mu$. These assumptions define the DI model:

$$\begin{aligned} \frac{dS}{dt} &= \mu(S^0 - S) - \lambda S, \\ \frac{dI_i}{dt} &= p_i \lambda S - (\mu + \nu_i) I_i, \qquad i = 1, \cdots, n, \\ \frac{dA}{dt} &= \sum_{j=1}^{n} \nu_j I_j - \delta A. \end{aligned} \tag{2.1}$$

The rate of infection, $\lambda$, depends upon the transmission probability per partner, $\beta_i$, of individuals in subgroup $i$, the proportion of individuals in the subgroup, $I_i/N$, and the number of partners of an individual per unit time, $r$. Simple random mixing leads to

$$\begin{aligned} \lambda(t) &= \sum_{i=1}^{n} \lambda_i(t), \\ \lambda_i(t) &= r\beta_i \frac{I_i(t)}{N(t)}, \end{aligned} \tag{2.2}$$

where $N = S + \sum_{j=1}^{n} I_j$.

Since we wish to examine the relative importance of each infection group in maintaining the chain of transmission, we will be concerned with the relative fraction of individuals being infected by each group, which we call the *relative impact* of the group. This fraction is

$$\rho_i(t) = \frac{\lambda_i(t)}{\lambda(t)} = \frac{\beta_i I_i}{\sum_{j=1}^{n} \beta_j I_j}. \tag{2.3}$$

## 2.2 Mathematical Analysis of the Model

Because transmission by individuals in group $A$ has been neglected under our assumptions, the transmission dynamics of (2.1) are determined by the transmission dynamics of the first two equations in system (2.1) and equation (2.2).

### 2.2.1 The Reproductive Number

The infected subgroups are linked in such a way that one infected subgroup cannot go to zero unless all of the infected subgroups go to zero. Therefore, system (2.1) has only two kinds of equilibria: the infection-free equilibrium given by $(S = S^0, I_i = 0)$ and the endemic equilibrium given by $(S = S^* > 0, I_i = I_i^* > 0)$. Analyzing the stability of the infection-free equilibrium gives the epidemic threshold condition, which specifies the conditions under which the number of HIV-infected individuals will increase when there are a small number of them or will decrease to zero otherwise. A simple stability analysis of (2.1), done by linearizing around the infection-free equilibrium and determining when the largest real part of the eigenvalues crosses zero, gives the threshold condition, characterized by the reproductive number

$$R_0 = r \sum_{i=1}^{n} \frac{p_i \beta_i}{\mu + \nu_i}. \tag{2.4}$$

If $R_0 < 1$, the infection-free equilibrium is locally asymptotically stable. If $R_0 > 1$, the infection-free equilibrium is unstable, and an initial infection will spread. The proof is given in Appendix A.

We can rewrite the reproductive number in a more intuitive and useful way as the product of the mean duration of infection, the average number of partners per unit time, and the mean probability of transmission per partner. This form for $R_0$ holds for all of the models studied here and allows to it be reinterpreted as the average number of individuals that a single infected individual will infect in a naive population.

For the DI model, the mean duration of infectiousness of an infected individual in group $i$ is $1/(\mu+\nu_i)$. Because $p_i$ of the infected individuals enter group $i$, the mean duration of infectiousness for all infected individuals in this model is

$$\bar{\tau}^D = \sum_{i=1}^{n} \frac{p_i}{\mu+\nu_i}. \tag{2.5}$$

There are several ways that we could define the mean probability of transmission from an infected individual in the population. We define it so that the average number of partners per unit time, $r$, times the mean probability of transmission gives the average number of individuals an infected individual will infect throughout the course of infection, given that none of his partners are infected before he has contact with them. Thus, the probability of transmission is weighted by the duration of infection,

$$\bar{\beta}^D = \frac{1}{\bar{\tau}^D} \sum_{i=1}^{n} \frac{p_i \beta_i}{\mu+\nu_i}. \tag{2.6}$$

These definitions give

$$R_0^D = r\bar{\beta}^D\bar{\tau}^D.$$

### 2.2.2 Endemic Equilibrium

To define an explicit formula for the endemic equilibrium, $(S^*, I^*)$, when $R_0 > 1$ we set the right-hand sides of (2.1) equal to zero. Then,

$$\mu(S^0 - S^*) = \lambda S^* = (\mu+\nu_i)I_i^*/p_i, \tag{2.7}$$

which gives

$$I_i^* = \mu(S^0 - S^*)\frac{p_i}{(\mu+\nu_i)}, \tag{2.8}$$

and then at the endemic equilibrium, the rate of infection is

$$\lambda^* = r\sum_{i=1}^{n}\beta_i\frac{I_j^*}{N^*} = r\sum_{i=1}^{n}\beta_i\frac{\mu(S^0-S^*)p_i}{(\mu+\nu_i)N^*} = \frac{\mu(S^0-S^*)R_0}{N^*}. \tag{2.9}$$

Also, it follows from (2.7) that

$$\lambda^* = \frac{\mu(S^0-S^*)}{S^*}. \tag{2.10}$$

As both (2.10) and (2.9) hold, $N^* = R_0S^*$.

Denote the total number of infected individuals by $I^{tot*}$. Then, $I^{tot*} = S^*(R_0-1)$. Using (2.5), we also have

$$I^{tot*} = \sum_{i=1}^{n} I_i^* = \mu(S^0-S^*)\bar{\tau}^D,$$

and hence

$$S^* + \mu(S^0-S^*)\bar{\tau}^D = R_0S^*. \tag{2.11}$$

Solving (2.11) for $S^*$, we arrive at

$$S^* = \frac{\mu\bar{\tau}^D S^0}{\mu\bar{\tau}^D + R_0 - 1}. \tag{2.12}$$

Combining (2.8) and (2.12), we have

$$I_i^* = \frac{\mu S^0(R_0-1)p_i}{(\mu\bar{\tau}^D + R_0 - 1)(\mu+\nu_i)} = \frac{p_i(R_0-1)S^*}{\bar{\tau}^D(\mu+\nu_i)}. \tag{2.13}$$

From (2.10) and the expression for $S^*$,

$$\lambda^* = \frac{\mu}{S^*}\left(S^0 - \frac{\mu\bar{\tau}^D S^0}{\mu\bar{\tau}^D + R_0 - 1}\right) = \frac{\mu S^0(R_0-1)}{S^*(\mu\bar{\tau}^D + R_0 - 1)} = \frac{R_0-1}{\bar{\tau}^D}. \tag{2.14}$$

It follows from (2.11) that we can rewrite the susceptible population as

$$S^* = \frac{\mu S^0}{\mu + \lambda^*}.$$

All components are positive at the endemic equilibrium if and only if the reproductive number $R_0$ is greater than 1. Moreover, the endemic equilibrium is always locally asymptotically stable whenever it exists, i. e., when it lies in physical space with all positive components. (The detailed proof is given in Appendix B.) In summary, we have the following theorem:

**Theorem 2.1.** *There exists a nonzero equilibrium given by* (2.12) *and* (2.13) *if and only if the reproductive number* $R_0$ *is greater than 1. If this endemic equilibrium exists, it is always locally asymptotically stable.*

Substituting the equilibrium formula for $I_i$ into the formula for the relative impact, we get

$$\begin{aligned}\rho_i^* &= \frac{p_i\beta_i}{(\mu+\nu_i)\sum\limits_{j=1}^{n} p_j\beta_j/(\mu+\nu_j)} \\ &= \frac{rp_i\beta_i}{(\mu+\nu_i)R_0} = \frac{p_i\beta_i}{(\mu+\nu_i)\bar{\beta}\bar{\tau}}.\end{aligned} \tag{2.15}$$

Note that at equilibrium, the fraction of infecteds in each group is independent of the per partner contact rate, but numerical studies [32] show that this is not true early in the epidemic. The larger the contact rate, the more important the most infectious group is to the early epidemic.

# 3 The SP Model

As in the DI model above, we assume that the susceptible population is homogeneous and is maintained by the same type of inflow and outflow. Assume that the population of infected individuals are subdivided into subgroups $I_1, I_2, \cdots, I_n$ with different infection stages such that infected susceptible individuals enter the first subgroup $I_1$ and then gradually progress from subgroup $I_1$ finally to subgroup $I_n$. Let $\gamma_i$ be the average rate of progression from subgroup $i$ to subgroup $i+1$, for $i = 1, \cdots, n-1$, and let $\gamma_n$ be the rate at which infected individuals in subgroup $I_n$ become sexually inactive or uninfectious due to end-stage disease or behavior changes. Then the dynamics of the transmission are governed by the following SP model:

$$\begin{aligned}\frac{dS}{dt} &= \mu(S^0 - S) - \lambda S, \\ \frac{dI_1}{dt} &= \lambda S - (\gamma_1 + \mu)I_1, \\ \frac{dI_i}{dt} &= \gamma_{i-1}I_{i-1} - (\gamma_i + \mu)I_i, \qquad 2 \leq i \leq n, \\ \frac{dA}{dt} &= \gamma_n I_n - \delta A,\end{aligned} \tag{3.1}$$

where the infection rate $\lambda$ is given by

$$\begin{aligned}\lambda &= \sum_{i=1}^{n} \lambda_i \\ \lambda_i &= r\beta_i \frac{I_i}{N}.\end{aligned} \tag{3.2}$$

Here, $r$ is the average number of partners per individual per unit of time, $\beta_i$ is the transmission probability per partner with an infected individual in subgroup $i$, $\delta \geq \mu$, and $N = S + \sum_{i=1}^{n} I_i$. Notice that the transmission by the $A$ group is neglected just as it was in the DI model. We also define the relative impact of the group as the fraction of individuals being infected whose infecting partner comes from group $i$:

$$\rho_i = \frac{\lambda_i}{\lambda}. \tag{3.3}$$

Like the DI model, the SP model has two equilibria: the infection-free equilibrium and the positive endemic equilibrium.

By investigating the local stability of the infection-free equilibrium, a straightforward calculation shows that the reproductive number can be defined by

$$R_0^S = r \sum_{k=1}^{n} \frac{\beta_k q_k}{\gamma_k + \mu}, \tag{3.4}$$

where we define

$$q_i := \prod_{j=1}^{i-1} \frac{\gamma_j}{\mu + \gamma_j}. \tag{3.5}$$

When $R_0 > 1$, the infection-free equilibrium is unstable, and thus the number of infected individuals will grow when a small number of individuals are infected. The epidemic will die out in the neighborhood of the infection-free equilibrium when $R_0 < 1$.

The mean duration of infection for the SP model is

$$\bar{\tau}^S = \sum_{i=1}^{n} \frac{q_i}{\mu + \gamma_i}, \tag{3.6}$$

where $1/(\mu + \gamma_j)$ is the average time period that infected individuals, who survive to infection stage $j+1$, spend in infection stage $j$, or the death-adjusted expected time in stage $j$ [23]. Because $1/\gamma_j$ is the waiting time in stage $j$ [18] (i.e., the mean time that an individual who progresses to stage $j+1$ spends in stage $j$), $\gamma_j/(\mu+\gamma_j)$ is the probability that an infected individual with infection stage $j$ survives to stage $j+1$, and $q_i$ is the total probability that an infected individual survives to stage $i$.

The mean probability of transmission per partner from an individual during the course of infection is

$$\bar{\beta}^S = \sum_{i=1}^{n} \beta_i \cdot (\text{fraction of time spent in stage } i).$$

The fraction of time spent in stage $i$ during the course of infection is the probability of reaching stage $i$ times the mean time spent in stage $i$ once it is reached, all divided by the mean duration of infection. The probability of entering stage $i$ is the probability of entering the previous stage, $i-1$, times the probability $\gamma_{i-1}/(\gamma_{i-1}+\mu)$ of entering stage $i$, given that the individual has entered stage $i-1$. Thus, the probability that an infected individual reaches stage $i$ is $\prod_{j=1}^{i-1} \frac{\gamma_j}{\gamma_j + \mu}$, or $q_i$. Because the mean time spent in stage $i$ is $1/(\gamma_i + \mu)$, we obtain

$$\bar{\beta}^S = \frac{1}{\bar{\tau}^S} \sum_{i=1}^{n} \frac{\beta_i q_i}{\gamma_i + \mu}. \tag{3.7}$$

These definitions allow us to rewrite the reproductive number formula for the SP model in the same form as that of the DI model:

$$R_0^S = r\bar{\beta}^S\bar{\tau}^S,$$

or, in other words, it is the average number of individuals an infected individual will infect early in the epidemic when none of his or her partners are infected by someone else.

In Appendix C we show that the endemic equilibrium for the SP model can be explicitly expressed by

$$\begin{aligned} S^* &= \frac{\mu \bar{\tau}^S S^0}{\mu \bar{\tau}^S + R_0^S - 1}, \\ I_i^* &= \frac{S^*(R_0^S - 1) q_i}{\bar{\tau}^S(\gamma_i + \mu)}. \end{aligned} \tag{3.8}$$

It follows from (3.8) that there exists a unique endemic equilibrium if and only if $R_0 > 1$.

Then, the total number of infected individuals is

$$I^{tot*} = \sum_{i=1}^{n} I_i^* = S^*(R_0^S - 1),$$

and

$$\lambda^* = \frac{r}{N^*} \sum_{i=1}^{n} \beta_i I_i^* = \frac{r(R_0^S - 1)}{R_0^S \bar{\tau}^S} \sum_{i=1}^{n} \frac{\beta_i q_i}{\gamma_i + \mu} = \frac{R_0^S - 1}{\bar{\tau}^S}. \tag{3.9}$$

Finally, we wish to know the fraction infected by each group once equilibrium is reached:

$$\rho_i^* = \frac{r \beta_i I_i \bar{\tau}^S}{S^* R_0^S (R_0^S - 1)} = \frac{r \beta_i q_i}{(\mu + \gamma_i) R_0^S}. \tag{3.10}$$

Table 1 consolidates our results for both DI and SP models.

Note that all of these formulas have the same form, with $p_i$ and $\nu_i$ from the DI model being replaced by $q_i$ and $\gamma_i$ for the SP model formulas. However, while it could be argued that $\nu_i$ and $\gamma_i$ are both progression rates and thus play somewhat similar roles in both models, $q_i$ is quite different from $p_i$. Not only is $q_i$ a derivative quantity, but $q_1 = 1$, so that the sum of the $q_i$ is larger than one, whereas the $p_i$ sum to one. Thus we should not let the similarity of form fool us into thinking the two models are the same.

Table 1: Formulas for the Two Models

| Name | DI Model | SP Model | Name | DI Model | SP Model |
|---|---|---|---|---|---|
| $R_0$ | $r\bar{\tau}\bar{\beta}$ | $r\bar{\tau}\bar{\beta}$ | $S^*$ | $\dfrac{\mu S^0}{\mu+\lambda^*}$ | $\dfrac{\mu S^0}{\mu+\lambda^*}$ |
| $\bar{\tau}$ | $\displaystyle\sum_{i=1}^{n}\frac{p_i}{\mu+\nu_i}$ | $\displaystyle\sum_{i=1}^{n}\frac{q_i}{\mu+\gamma_i}$ | $I_i^*$ | $\dfrac{p_i S^*\lambda^*}{\mu+\nu_i}$ | $\dfrac{q_i S^*\lambda^*}{\mu+\gamma_i}$ |
| $\bar{\beta}$ | $\displaystyle\sum_{i=1}^{n}\frac{p_i\beta_i}{\bar{\tau}(\mu+\nu_i)}$ | $\displaystyle\sum_{i=1}^{n}\frac{q_i\beta_i}{\bar{\tau}(\mu+\gamma_i)}$ | $I^{tot*}$ | $S^*(R_0-1)$ | $S^*(R_0-1)$ |
| $\lambda^*$ | $\dfrac{R_0^D-1}{\bar{\tau}^D}$ | $\dfrac{R_0^S-1}{\bar{\tau}^S}$ | $\rho_i^*$ | $\dfrac{p_i\beta_i}{(\mu+\nu_i)\bar{\beta}\bar{\tau}}$ | $\dfrac{q_i\beta_i}{(\mu+\gamma_i)\bar{\beta}\bar{\tau}}$ |

# 4 Model Simulations

## 4.1 Parameter Estimation

Here we review studies and data on the parameters for both the DI and SP models and obtain estimates for the numerical simulations in the next section. Many of these parameters have wide ranges of uncertainty. We choose a baseline set of parameters that lies in the center of this range. In this paper we examine the sensitivity of the models to one of our parameters, the probability of transmission per contact: in [32] we will present sensitivity studies for the other parameters.

### 4.1.1 Parameters Common to Both Models

**Natural Death Rate** We split up the removal rate, $\mu$, into the natural death rate, $d$, and the rate, $\alpha$, at which individuals leave the high-risk population due to migration and changes in sexual behavior. Thus $\mu = d + \alpha$. We assume that individuals in our population are young adults and can expect to live an average of 50 more years. Thus $d = 0.02$ yrs$^{-1}$.

**Mean Time in the High-Risk Group** The number of years that people engage in high-risk behavior is unknown and probably varies greatly between populations.

We take a baseline of 20 years. In [32] we will study how changing $\alpha^{-1}$ from 5 to 50 years affects our model results. Note that increasing $\alpha$ decreases the mean time that an individual spends in the sexually active population in different ways for each model. Thus it implies that the two models will have different reproductive numbers as soon as $\alpha \neq 0$ unless we modify other parameters to hold $\bar{\tau}$ constant.

**Mean Duration of Infection** One of the best statistical analyses of progression to AIDS is that in [18], where the authors used a staged-progression model and found a mean time from infection to AIDS of 8.6 years. However, this study was done in 1989 and thus did not have access to data on long-term progressors, since the epidemic had not been around long enough, and could not take into account the impact of treatments that have been developed since 1989. A careful statistical analysis of the duration of infection is outside of the scope of this paper, but we look at data from two recent papers [4, 33] and make rough estimates of their progression rates. All of our estimates came out longer than 8.6 years, since they included data on longer-term progressors than any people included in [18]. This increase is consistent with past trends: statistical estimates for the mean duration of infection have tended to increase as data has accumulated on long-term survivors, and new treatments have increased life expectancies.

In a recently published study on HIV-RNA levels [4], the authors present Kaplan-Meier curves for times from infection to AIDS for each of their four groups. By reading numbers off of their curves and weighting them by the fraction of the population in each group, we obtained a crude estimate of 19 years from infection to AIDS. This estimate is so much longer than that in [18] that we didn't use it. However, we did use the relative progression rates for each group in [4] for the DI model (see below).

In a more recent study [33], the population was divided according to HIV RNA plasma levels. Most participants were already infected at entry, and the start point used in their analysis was either 1 or 1.5 years after entry (the time when the first reliable plasma sample was taken). The authors provided a table of the fraction of participants with AIDS at 3, 6, and 9 years after the start date. This study found a similar increase to that in [4] in progression rates with each increase in RNA levels. Those progression rates almost doubled for each factor-of-three increase in RNA levels. We are assuming that people were infected 1.5 years before entering the study, estimate $\nu_i$ for each RNA group separately, and add in the natural death rate to obtain a mean time from infection to AIDS of between 9 and 13.5 years for this population, depending on which of the data points we use (of the 3, 6, and 9 years). This lies between the 8.6 years of [18] and the 19 years of [33]. For our baseline value, we take $\bar{\tau} = 12$ years, which is in the middle of all of these estimates. Note that this estimate assumes that $\alpha = 0$, since these studies keep people in the cohort, regardless of their sexual behavior.

Looking at the Kaplan-Meier curves shows one reason for the large variation in our estimates from those in [33]: the population is not developing AIDS in an exponential manner. Instead, progression rates are very small early in infection and gradually increase. Our staged progression model can account for this, but the DI model cannot.

**Transmission Probability and Partner Acquisition Rate** The probability of transmission per partner depends upon the average number of contacts per partner and the mean probability of transmission per contact.

Although studies on sexual behavior are difficult, it appears that partner acquisition rates vary a great deal between communities and that HIV can spread even in populations with fairly low values of $r$. Highly active subpopulations of homosexual men and prostitutes have reported hundreds of sex partners per year, whereas in African populations in which HIV has spread, reported partner acquisition rates may be as low as 1 or 2 per year. We take our baseline value of $r$ to be 5 partners per year, since this is typical of many populations in which HIV spreads. However, since $r$ can vary so much, in [32] we check the sensitivity of our model results to values of $r$ ranging from 1 to 100 partners per year.

We use $z$ as the transmission probability per sexual contact. Estimates of $z$ range from 0.0003 (lowest value estimated for female to male) to 0.08 (highest value estimated for male-to-male transmission) [15]. This range is large, and our models are very sensitive to this parameter. We take $z = 0.003$ as our baseline value, partly because it is close to the value found in a number of couples studies, partly because it lies on the low side of the possible range for $z$, and partly because the reproductive number ends up near 2, given all of our other baseline choices.

If $n$ is the number of contacts with a partner, then the average probability of transmission per partner is $\bar{\beta} = 1 - (1-z)^n$. There is no known relationship between the average number of contacts per partner and the number of partners per year. In general, the number of contacts per partner will decrease as the partner acquisition rate, $r$, increases. We choose a simple function with this property to be $n = 104/r+1$. Then $\bar{\beta} = 1 - (1-z)^{104/r+1}$. We are assuming that people with few partners have 2 contacts per week and people with many partners have slightly more than one contact per partner. At the baseline value of $r$, this gives a mean of 21.8 contacts per partner, and $\bar{\beta} = 0.063$.

Since $z$ is in the interval $(0.0003, 0.08)$, in a population with 30 partners per year, $\bar{\beta}$ lies in the range 0.0013 to 0.31, and the reproductive number for both models ranges from 0.48 to 112 when $\bar{\tau} = 12$ years. In a population with 3 partners per year, $\bar{\beta}$ lies between 0.011 and 0.95, giving a reproductive number of 0.38 to 51. Notice how the variability in estimates for the probability of transmission leads to such a large uncertainty in the reproductive number and, thus, in the dynamics of the epidemic, while the dependence of $\bar{\beta}$ on $r$ prevents $r$ having such a large effect on the epidemic, especially at lower values of $z$. If we make a different assumption about how the number of contacts per partner varies with $r$, then, of course, we would get more or less sensitivity of our results to $r$. In fact, in [32] we show that if $n$ drops more rapidly as $r$ increases, the reproductive number can be nonmonotonic.

In both scenarios, if the transmission probability is actually at its lowest estimated value, then the epidemic will not spread, whereas at its highest value, the epidemic will spread very rapidly and be very difficult to stop. This is one reason why condom usage, which can reduce the probability of transmission per contact by 90% or more, can have such a dramatic effect on the spreading of the epidemic.

**Infectivity of the Subgroups** Our models, and the theoretical results we derived from them, are not dependent on the assumption that RNA viral levels in serum and transmissibility of the virus are correlated. However, we now make this assumption in order to derive relative values for the subgroups in each model to use in our numerical studies. Although it seems highly plausible that viral levels in serum and plasma are correlated with infectiousness, it is important to keep in mind that such a correlation does not necessarily exist. Assuming that it does assumes that serum levels are correlated with membrane and secretion levels, and that those in turn somehow determine transmissibility itself. Newly developed lab techniques allow researchers to begin to measure HIV levels in semen, membranes, and cervical and vaginal fluids and to examine whether or not these levels correlate with plasma and serum levels. As pointed out by Royce et al. [15] in their review of sexual transmission, results on this correlation to date are inconsistent: some studies, such as those in [34, 35, 36], have found a correlation; others, such as those in [37], have not.

There is other evidence that HIV serum and plasma levels or individual variations affect transmission, but it is less direct. Maternal HIV-1 RNA levels affect the probability of transmission to the fetus [38]. Occasional reports of superspreaders (individuals who have infected large numbers of partners) (see [9, 39], and the discussion in [10]) or the odd case of a dentist apparently infecting several patients [40] provide additional support for the hypothesis that some individuals are highly infectious over long periods of time. It seems reasonable to assume that genetic factors can partially determine an infected individual's infectiousness, especially since they are known to partly determine susceptibility. Other factors, such as viral strain, age (which affects progression rates [4, 41]), the presence of other sexually transmitted diseases [42], smoking [43], general health, individual chemistry, and pregnancy [44], may affect an individual's ability to transmit HIV.

### 4.1.2 Parameters Specific to the DI Model

The remaining parameters for the DI model are obtained from the HIV progression study reported in [4], where the authors measured HIV RNA levels in a long-term study of hemophiliacs. They used serum samples that had been archived between 12 and 36 months after their estimated seroconversion dates. This time interval was chosen in order to measure levels during the chronic HIV RNA phase: after 12 months the initial pre-antibody peak was over, and prior to 36 months the increase in levels at the end of the chronic phase had generally not begun.

O'Brien et al. [4] also divided the subjects according to HIV levels. Nineteen individuals had 0-999 copies/ml, eighty-two had 1,000-9,999 copies/ml, fifty-five had 10,000-99,999 copies/ml, and nine had 100,000+ copies/ml. Of the last group, one person had more than a million copies/ml. They presented Kaplan-Meier survival curves for these 4 groups. Survival patterns were dramatically different for each of the groups, with the group having the highest viral load dying the most quickly. Assuming exponential decay for each population, crudely estimating the decay rate from their graph showed that the death rate nearly doubled for each factor-of-10 increase in viral loads.

O'Brien et al. [4] then argued that viral levels fluctuated very little within an individual during the chronic phase in the absence of treatment (which can decrease viral levels by several orders of magnitude). For 62 of their 165 subjects, more than one sample was available. The difference in HIV-1 RNA levels between specimens was less than a factor of 10 for 53 of the subjects and less than a factor of 100 for the remaining 9 individuals. Of course, this also says that the picture of constant viral load is an oversimplification: viral loads fluctuate some in all individuals [6].

Following the study in [4], we divide the infected population into four groups ranging from the highest viral load to the lowest. We assume a linear connection between infectivity and viral load so that each group's infectivity decreases by a factor of 10 (and study the sensitivity of the model to this assumption). Rounding these numbers, we yield

$$\mathbf{p} = (0.05, 0.33, 0.5, 0.12)^T, \quad \boldsymbol{\beta}^D = (1000, 100, 10, 1)^T b^D,$$

where the scalar parameter $b^D$ is to be chosen so that the mean probability of transmission is what we want it to be. Here we have used matrix transpose notation to give the numbers.

When we used the Kaplan-Meier curves from [4] to estimate progression rates for each of their four groups, we obtain $\boldsymbol{\nu} = (0.1, 0.05, 0.03, 015)^T$ yrs$^{-1}$. This gives a mean duration of infection of 18.6 years. To modify this to give the mean of 12 years that we have assumed, we increase each $\nu_i$ by a factor of 1.92 so that

$$\boldsymbol{\nu} = (0.19, 0.096, 0.058, 0.028)^T \text{ yrs}^{-1}$$

when these baseline parameters are substituted into equation 2.5 with $\alpha = 0$. Note that as soon as $\alpha > 0$, $\bar{\tau}^D$ is less than 12 years. Substituting all of the above parameter choices into equation 2.6 gives a baseline mean probability of transmission per partner (with $\alpha = 0.05$ yrs$^{-1}$) of

$$\bar{\beta} = 59.0 b^D.$$

Specifying $\bar{\beta}$ then allows us to calculate the appropriate value for $b^D$.

### 4.1.3 Parameters Specific to the SP Model

For the SP model, we assume that the population goes through 4 stages: an early, highly infectious pre-antibody phase, two chronic stages at low infectiousness, and a final stage at higher infectiousness. According to [2, 6, 21], the peak in viral loads occurs 2 to 6 weeks after infection in the majority of patients, after which viral levels decline rapidly over the next 1 to 2 weeks. Viral loads at their peak may be as high as $10^6$ or $10^7$ copies/ml. From the data presented above, we know that after this early phase viral loads are rarely this large. In fact, the majority of infected individuals have between $10^3$ and $10^4$ copies/ml. Late in infection, viral loads may rise but usually by less than tenfold. Therapy tends to reduce viral levels, but often only temporarily.

We need to choose our parameters so that the mean duration of infection is 12 years when $\alpha = 0$. If we take 4 weeks as the duration of the initial stage, 3 years as the mean duration of the final and more infectious stage, and assume the middle two stages to have an equal duration, with a rate of moving on to the next stage denoted by $\gamma_m$, then

$$\gamma_1 = 13 \text{ yrs}^{-1}, \;\gamma_2 = \gamma_3 = \gamma_m, \;\gamma_4 = 0.333 \text{ yrs}^{-1}.$$

Substituting these values into (3.5) for $q_i$, setting $\bar{\tau} = 12$ yrs, and then solving (3.6) gives

$$\gamma_m = 0.177 \text{ yrs}^{-1}.$$

Thus each of the middle stages lasts an average of 5.7 years.

Given the data on viral loads, we assume for the SP model that

$$\boldsymbol{\beta}^S = (100, 1, 1, 10)^T b^S,$$

where $b^S$ is a scalar parameter that determines the mean probability of transmission. Using the $\gamma_i$ obtained above, we get $\bar{\beta} = 3.345 b^S$.

#### 4.1.4 Summary

We have chosen the baseline parameters for each model to represent our best estimate for fitting the model to the current HIV epidemic in the United States. These parameters, defined in Table 2, are selected based on the analysis of epidemiological data presented in Sections 4.1.1–4.1.3. For both the DI and SP models, the mean duration of infection is $\bar{\tau} = 12$ yrs, which is obtained by setting $\alpha = 0$. However, the mean time an infected person stays in the active population ($\alpha = 0.05$) is shorter in the DI model ($\hat{\tau}^D = 7.3$ yrs ) than in the SP ($\hat{\tau}^S = 8.26$ yrs). Therefore, the reproductive number for the DI model is smaller than for the SP model.

### 4.2 Numerical Simulations

Our goal in these examples is to investigate the relative impact of the infected subgroups in the DI and SP models on the epidemic. Although the two models were derived using very different assumptions about the biology of infected individuals, if the mean duration of infection and the mean infectivity are the same for the two models, then their reproductive numbers and the numbers of total infected individuals at the endemic equilibrium will be the same. This implies that, without direct evidence

Table 2: Baseline Parameters

| | | |
|---|---|---|
| Initial population size | $N = S(0) + I_{tot}(0)$ | 1.0 |
| Initial infected population | $I_{tot}(0)$ | 0.01 |
| Natural death rate | $d$ | 0.02 $\text{yrs}^{-1}$ |
| Sexually active removal rate | $\alpha$ | 0.05 $\text{yrs}^{-1}$ |
| Total removal rate | $\mu = d + \alpha$ | 0.07 $\text{yrs}^{-1}$ |
| Mean duration of infection ($\alpha = 0$) | $\bar{\tau}$ | 12 years |
| Partner acquisition rate | $r$ | 5 partners/year |
| Mean probability of transmission per contact | $z$ | 0.003 |
| Mean probability of transmission per partner | $\bar{\beta}$ | 0.063 |
| **DI parameters** | | |
| Distribution by group upon infection | $p_i$ | (0.05, 0.33, 0.5, 0.12) |
| Progression rates by group | $\nu_i$ | (0.19, 0.096, 0.058, 0.028) |
| Relative infection rates | $\beta_i$ | $(10^3, 10^2, 10, 1)b^D$ |
| Mean duration of infectivity | $\hat{\tau}^D$ | 7.3 yrs |
| Reproductive number | $R_0$ | 2.3 |
| | $b^D = \bar{\beta}/59.0$ | 0.00107 |
| **SP parameters** | | |
| Progression rates by group | $\gamma_i$ | $(13.0, 0.177, 0.177, 0.333)$ |
| Relative infection rates | $\beta_i$ | $(100, 1, 1, 10)b^S$ |
| Mean duration of infectivity | $\hat{\tau}^S$ | 8.26 yrs |
| Reproductive number | $R_0$ | 2.6 |
| | $b^S = \bar{\beta}/3.3$ | 0.019 |

of the manner in which infectivity varies between individuals, it is not possible to use these two quantities to tell which hypothesis is valid. However, even when the reproductive numbers and endemic equilibrium are the same, the transient behavior and internal dynamics of the models can be very different.

An added complication in any model with infected subgroups is the sensitivity of the model to the distribution of the initial infected population between the different subgroups. Different distributions of even a small infected population can hasten or delay the onset of an epidemic by several years. We approximate the natural initial conditions for an epidemic by introducing a very small infected group (0.01%) into the population and allowing the epidemic to progress until 1% of the population is infected. At this time, we re-normalize the population while maintaining the same relative distribution of the infected population and call this time $t = 0$. With this approach, we find the model to be insensitive to the distribution of the initial 0.01% infected population. We explore this issue in more detail in [32].

In all the examples, unless we explicitly state otherwise, we use the baseline parameters in Table 2.

#### 4.2.1 Relative Impact of the Infected Groups on the Epidemic

When there are multiple infected subgroups, each infected subgroup has a different impact on the spread of the disease. In this example, we compare the relative impact of the infected groups on the epidemic and observe how quickly these rates converge to their asymptotic values. This example identifies which infected groups are driving the epidemic. The hope is that this information may eventually help guide intervention strategies to slow the epidemic. In particular, the analysis can be used to compare the sensitivity of the epidemic by targeting intervention strategies focused on identifying people in the most infectious group in the DI model and the group in the SP model with the acute infectious period.

We use the *relative impact* of each infected group, $\rho_i(t) = \lambda_i(t)/\lambda(t)$, to study how fraction of the infections attributed to each infected group is directly related to the assumptions about the length of time spent in each group and the infectivity. In the numerical simulations, we monitor $\rho_i(t)$ directly and study how quickly these ratios converge to their asymptotic values given by (2.15) and (3.10).

**DI Model:** The numerical simulation for the DI model in Figure 1 demonstrates how the epidemic converges to the endemic equilibrium. The plot of the functions $\rho_i(t)$ shows that in the early spreading of the epidemic, the very small but highly infectious group $I_1$ causes the bulk of the infections. However, as the epidemic progresses, these individuals develop AIDS more rapidly than the individuals in the other groups and they have less impact than the larger (but less infectious) group $I_2$ . Interestingly enough, even though group $I_3$ is the largest group and stays in the population much longer than the first two groups, the assumption that they are tenfold less infectious prevents this group from ever having a major impact on the epidemic. The moderate size and small infectiousness of group $I_4$ results in this long-lived group playing almost no role in the epidemic at any time.

**SP Model:** The solution for the SP model in Figure 2 confirms the observation made by Jacquez et al. [21, 30] that in the early epidemic almost all infections for the SP model are due to individuals in the first infection stage. Later on, this stage becomes less important until eventually the infections transmitted from people in the long lived last stage of the epidemic (Stage $I_4$) dominate. Stages $I_2$ and $I_3$ play less of a role throughout the epidemic.

**Comparison:** Although the macro dynamics of the total infected populations for the DI and SP are similar, the internal dynamics of the transmission process, as measured by the relative importance of the infected groups, are very different. These differences have significant implications if the models are used to gain insight into planning intervention strategies. For example, suppose that the most infectious group could be identified, perhaps through contact tracing, and convinced to change their behavior. In the DI model this would have an immediate and significant impact on the epidemic. In the early epidemic for the SP model, the most infectious group is

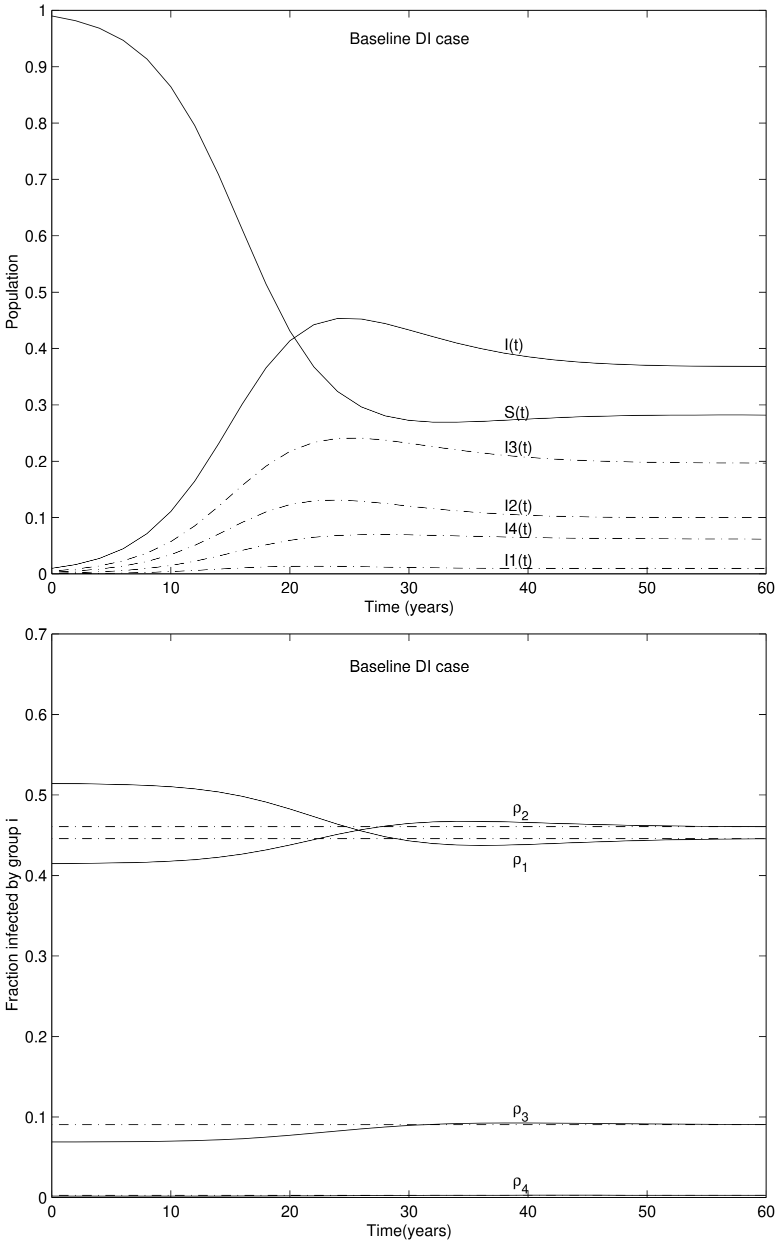


Figure 1: The solution and relative infection rates of the DI model for the baseline parameters (Table 2) show how quickly they converge to the equilibrium values (shown by the dashed lines). Note that the initial relative impact of the infected groups can be different from its asymptotic value. In fact, in this example the first infected group causes the largest number of infections early in the epidemic, and after about 25 years the second infected group becomes the most important in terms of transmission.

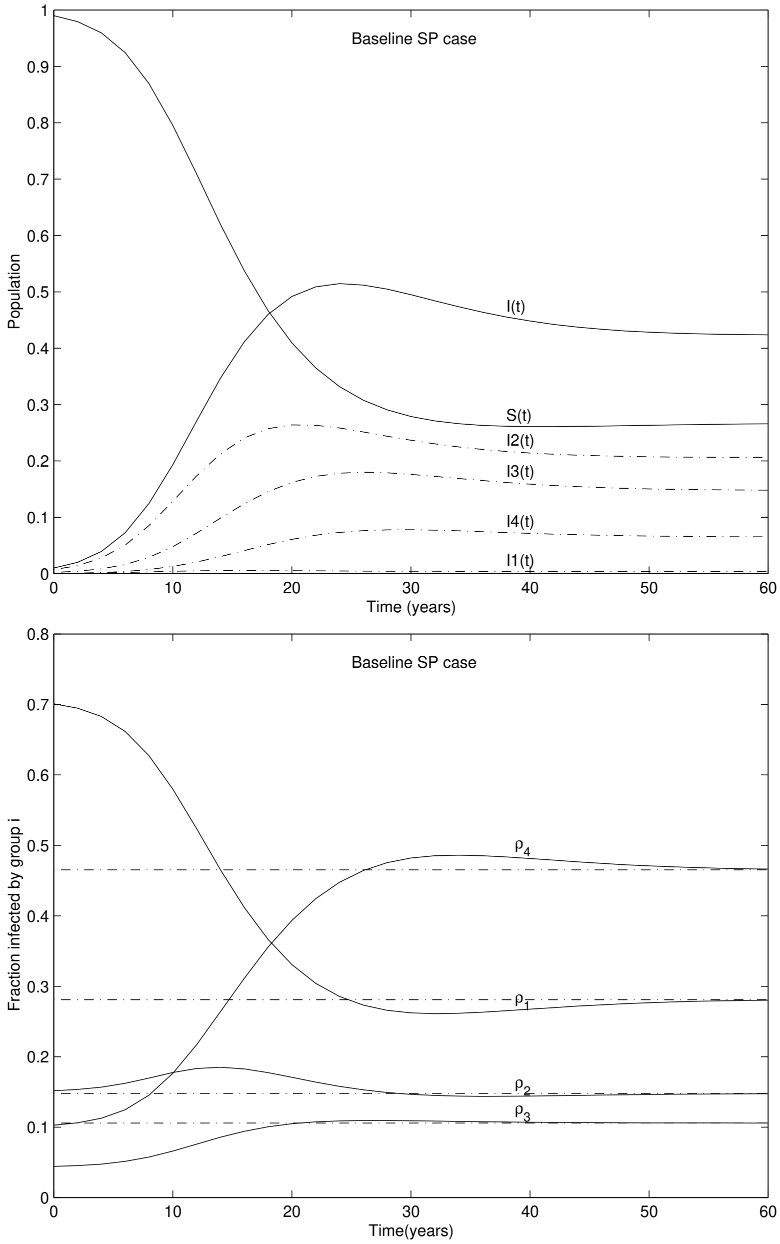


Figure 2: The solution and relative impacts of the SP model for the baseline case parameters given in Table 2. The plot of the populations shows the initial rise and subsequent convergence to the equilibrium values. Note that the initial relative impact of an infected group can be significantly different from its asymptotic value. In fact, in this example the relative impact of the first and fourth infected groups in the SP model switch places, in that early in the epidemic the $I_1$ group causes most of the infections, and later on the $I_4$ group becomes the most significant. Even though the other two groups are the largest, they transmit a small fraction of the infections.

the one in the first infection stage which is continually being replenished and the impact would be significantly less. On the other hand, if an approach could be used to identify more people in the very early stages of the infection, then this would have only a proportional effect on the DI model, but could have a large impact on infections caused by those in the first stage of the SP model. The insight gained by using a model to understand the impact of drug therapy to extend the life expectancy of infected individuals is also sensitive to which model is being used.

#### 4.2.2 Sensitivity to the Probability of Transmission Per Contact

When defining the baseline parameters, we made the assumption that the mean probability of transmission per partner is $\bar{\beta} = 1 - (1 - z)^{(104/r+1)}$, where $z$ is the mean probability of transmission per contact. Under this assumption, the reproductive number increases rapidly with $z$ when it is small, and gradually levels off. As pointed out above, $z$ is not well known. Not only are the ranges of $z$ from any study fairly large, but it also appears that it may vary greatly between populations. In this example, we demonstrate the sensitivity of both models to small changes in $z$.

**DI Model:** Figure 3 illustrates the extreme sensitivity of $R_0$ for the DI model as a function of $z$. At the minimum estimate for $z$, $R_0$ is below 1, and there is no epidemic. At the highest estimate of $z = 0.08$, $R_0$ is 31, and the epidemic is rapid and devastating. $R_0$ crosses 1 when $z = 0.00127$. When $R_0$ is near 1, there is a very small, slow epidemic, but as it increases, there is at first a rapid change in the behavior of the epidemic, which spreads more quickly and extensively.

The sensitivity can be further seen by the wide variations in the progression of the epidemic in Figure 4, where we compare $(z = 0.002, R_0 = 1.56)$, $(z = 0.003, R_0 = 2.31)$ and $(z = 0.004, R_0 = 3.06)$. Once $R_0$ is greater than four or five, the epidemics all progress similarly by rapidly infecting and depleting the susceptible population.

Even though $z$ has no impact on the equilibrium fractions of infections attributable to each group, as $z$ increases, group $I_1$ becomes relatively more important to the early epidemic than the other groups. Group $I_1$ contributes less and less as the

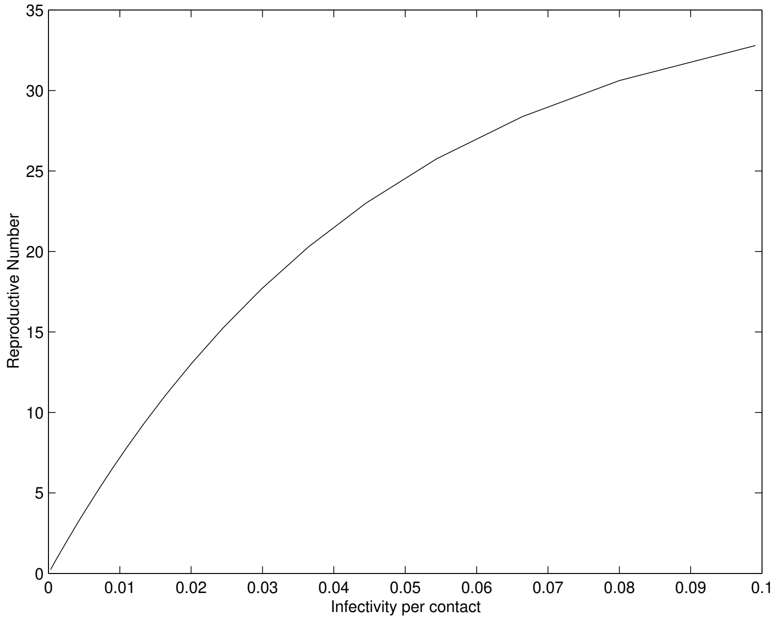


Figure 3: The reproductive number for the DI model is a sensitive function of the mean probability of transmission per contact $z$. All the other parameters are the baseline values given in Table 2.

epidemic reaches its peak. For these parameters, most of the infections at all times are attributable to groups $I_1$ and $I_2$. Although groups $I_3$ and $I_4$ make up the bulk of infected individuals, they have relatively little to do with spreading the epidemic

**SP Model:** Both the endemic states and the transient dynamics of the SP epidemic are sensitive to $z$. The susceptible population shown in Figure 5 at time $t = 20$ differs by a factor of four when the probability of transmission per contact is changed by a factor of two.

As in the DI model, the asymptotic fraction of the population infected by each group is independent of $z$, although the transient dynamics are sensitive to the partnership acquisition rates. When $z$ is large, the first (highly infectious) group $I_1$ progresses extremely rapidly and initially drives the epidemic. When $z$ is small, the early epidemic progresses much slower, and the last infected group $I_4$ is more important.

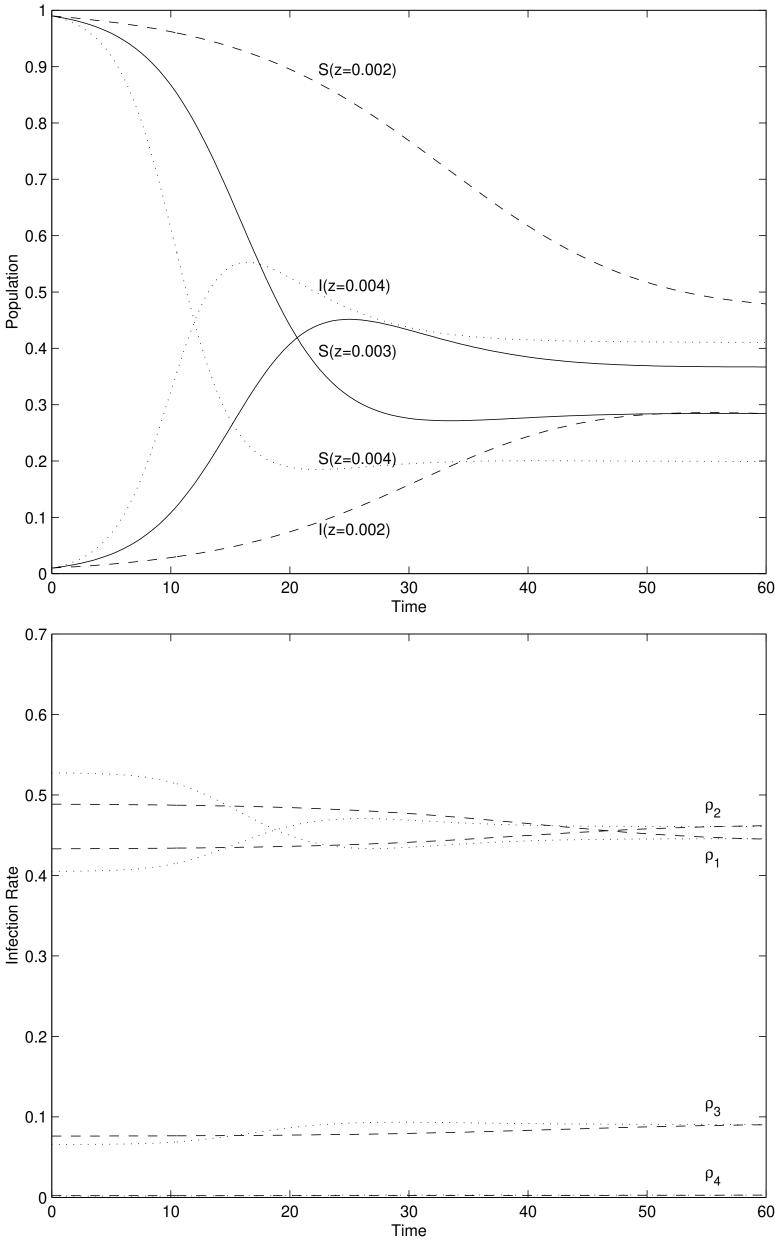


Figure 4: The solution of the DI model is extremely sensitive to the mean probability of transmission per contact. The baseline epidemic $z = 0.003$ (solid line) progresses more rapidly as $z$ increases and all parameters (except $b^D$) are held fixed as seen when $z = 0.002$, $b^D = 7.23\ 10^{-4}$ (dashed line), and when $z = 0.004$, $b^D = 1.42\ 10^{-3}$ (dotted line). The plot of the fractions infected shows the relative impact of the infected groups for the $z = 0.002$ (dashed line) and $z = 0.004$ (dotted line) cases

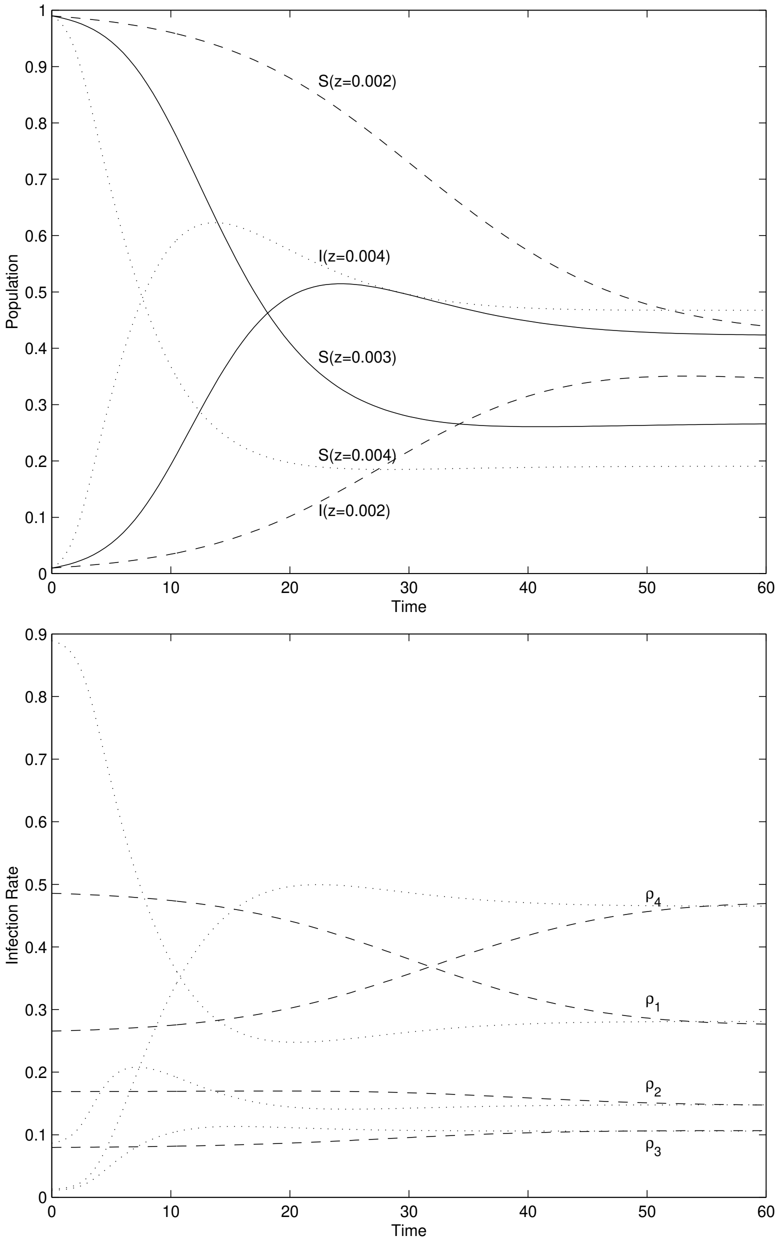


Figure 5: The SP model is also sensitive to the partnership acquisition rate $z$, even when the total number of contacts is fixed. In the population plot, the solid curves are the S and I populations for the baseline case of $z = 0.003$. The populations are also shown for the cases when $z = 0.002$ ,$b^S = 0.013$, $R_0 = 1.8$, (dashed line), and $z = 0.004$ ,$b^S = 0.025$, $R_0 = 3.4$ (dotted line). All of the other parameters are at the baseline values given in Table 1. In the infection rate plot note that the asymptotic fraction of population being infected by each group is independent of $z$. The relative importance of the different groups changes drastically during the epidemic when $z = 0.002$ (dashed line) and when $z = 0.004$ (dotted line), respectively.

**Comparison:** Both the DI and SP models are sensitive to the assumptions made for the probability of transmission per contact. In the absence of good data and sufficient model complexity, this sensitivity points out one reason it is difficult to use models for qualitative predictions of the AIDS epidemic. This is especially true in the early stages of the epidemic.

This sensitivity of the epidemic to the per-contact transmission probability demonstrates how important the use of condoms and spermicides are in preventing the spread of the epidemic. Decreasing $z$ may be able to drop the epidemic down below the threshold. If, as we have assumed, the number of contacts per partner increases as the number of partners decreases, changing $z$ can have more impact on the spread of HIV than changing the partner acquisition rate $r$.

We also investigate how the epidemic spreads more rapidly as the probability of transmission is increased. In the SP model, while the equilibrium fractions infected by each group are unaffected by changes in $z$, as the probability of transmission increases, the first group becomes even more important to the early epidemic. The same is true as the partner acquisition rate $r$ is increased ( plots not shown. See [32]) . Consistent with the modeling work in [21, 30], we observed that with an $r$ of 50, the first group causes more than 90% of all early infections, although its impact quickly drops to less than 30% at the endemic equilibrium. The impact of $r$ on who is causing the actual infections early in the epidemic is much greater for the SP model than for the DI model, where only a small change was seen even when $r$ was as large as 100.

# 5 Summary and Concluding Remarks

Based on the hypothesis that HIV-1 RNA levels measured in serum and plasma are correlated with infectiousness, two simple models were formulated and investigated, each capturing one of the observed aspects of variations in RNA levels. Our numerical studies use numbers which rely upon the somewhat tenuous association between HIV RNA levels in the bloodstream and infectiousness. Future studies are needed to establish this connection. It is also important to keep in mind that although new

RNA lab techniques are more reliable than older techniques, there remain questions about their accuracy [45].

The differential infectivity model, which has never been previously studied, accounts for differences between individuals, and the staged-progression model, which is similar to models previously studied, accounts for differences within the same individual over the course of infection. Although, undoubtedly, individuals vary in infectiousness both temporally and individually, and the most complete model should include a combination of these two hypotheses, it is insightful (and more mathematically tractable) to first consider them separately.

For both models, we derived explicit formulas for their reproductive numbers and endemic equilibria. These formulas were expressed in a similar and easily interpreted form for both models. For example, the reproductive number is $R_0 = r\bar{\beta}\bar{\tau}$ for both models. We showed that if the reproductive number is less than one, the infection-free equilibrium is the only equilibrium which is locally asymptotically stable. If the reproductive number is greater than one, the infection-free equilibrium becomes unstable, the epidemic spreads, and a unique endemic equilibrium appears. For the DI model we showed that this endemic equilibrium is locally asymptotically stable.

We defined a new quantity, the relative impact of the group, as the fraction of new infections being caused by that group. Then, using mid range parameters and estimates from cohort studies, we examined the transmission dynamics of these two models and the relative impact of each group. Despite very moderate choices for the partner acquisition rate and the transmission probability, both models had reproductive numbers greater than 2. We also showed that the reproductive number and many details of the epidemic are very sensitive to the transmission probability per contact.

For the DI model, and our parameter choices, the two most infectious groups are responsible for almost all of the transmissions, despite the fact that their life-expectancy is shorter than the other two groups. The most infectious group, entered by only 5% upon infection, is responsible for over 40% of all transmissions. For the SP model, those in the acute infectious phase transmit a very large fraction of the infections despite the very short duration of this phase. Most of the remaining infections are transmitted by those in the late chronic stage. These conclusions about who transmits infection are robust to changes in the probability of transmission per contact, despite the fact that the overall epidemic is very sensitive to $z$.

In the absence of a vaccine, epidemiologists must rely upon effecting behavior changes and providing condoms in order to control the spread of HIV. If infectious individuals could be identified and convinced to change their behavior, then perhaps the spread of HIV could be slowed or halted. If viral levels indicate infectiousness, then treatment of these individuals could also slow spreading by lowering their viral loads even when they are unwilling to change their behavior. Ideally, all infected individuals should be identified and provided with treatments. However, screening everyone is not possible, and HIV is an infection which remains asymptomatic for many years, so that most infected people do not know they are infected. Our study indicates that it is more urgent to identify some infected individuals than others.

The DI and SP models capture two different aspects of the HIV epidemic. More significantly, we believe that the two mechanisms imply different things about the best way to control the epidemic in the absence of a vaccine. In particular, contact tracing could be very effective at catching the most infectious individuals if a small group is responsible for most of the infections and if that group of individuals is very infectious throughout the chronic stage. Such superspreaders would have infected a number of partners, one of whom is likely to identify the superspreader in an interview, thus hopefully allowing the superspreader to be contacted and counseled and no longer spreading the virus to others. On the other hand, if every individual is infectious for only a short period of time at the beginning of the infection, then, by the time a person is named by someone who has been infected, that person will already be in the uninfectious chronic stage. In that case, awareness of the symptoms of early acute infection might be the best way to identify people before they infect others. Likewise, alertness to symptoms may be a fairly effective way, perhaps combined with general screening programs, to identify individuals before they enter a more infectious stage late in their infection.

There has not been space to present further examples. In a future paper [32] we will do extensive sensitivity studies of both models to all of the parameters and, in particular, to the migration rate and the relative infectivities of the different groups. We will also carefully address the neglected question of initial conditions for these types of models, showing that model results can be extremely sensitive to initial conditions. We will develop a "natural initialization procedure," that will be robust and will work for many population dynamics models.

It is important to explore and understand the DI and SP models separately before going to a model that has more complexity. However, the HIV RNA data indicate that a combined model might be necessary to capture some very important characteristics of the epidemic. This is indeed the case if we wish to settle the question of who causes most of the infections and, thus, whether or not contact tracing is a cost-effective way of controlling the HIV epidemic. In a future paper, we plan to explore a DISP model in which the infected population is divided into $n \times m$ groups, where $n$ is the number of different stages of infection and $m$ is the number of inherently different groups. Note that we have also neglected many other important features of the AIDS epidemic, such as variations in sexual behavior, age, and inherent susceptibility.

Although in this paper we have not explored the possibility that individuals vary inherently in their susceptibility, there is growing evidence to support this inherent variation. Some individuals appear to be genetically immune to infection [17], and others seem to be more susceptible [46]. Circumcision appears to decrease susceptibility in men [47], and the presence of other sexually transmitted diseases seems to increase susceptibility [14, 47]. Variations in susceptibility should not affect epidemic dynamics much until an epidemic is far advanced and the more susceptible individuals have been depleted from the population, unless susceptibility and infectiousness are linked.

It is clear from our preliminary explorations of these two models that most HIV infected individuals are not spreading the infection. If the subgroups that are transmitting infection could somehow be identified and convinced to refrain from risky behaviors, at least while they are infectious, the epidemic could perhaps be contained. Unfortunately, if the SP model results are to be believed, a lot of the spread occurs during the first few weeks to two months after infection. While a large number of infecteds have acute primary infection symptoms, many do not. Further study with a combined model will tell us whether this group is truly important or whether the superspreaders of the DI model are more important.

# References


[1] Pratt, R. D., J. F. Shapiro, N. McKinney, S. Kwok, and S. A. Spector, *Virologic characterization of primary HIV-1 infection in a health care worker following needlestick injury*, J Infect. Dis, 172 (1995), 851–854.

[2] Quinn, T. C., *Acute primary HIV infection*, JAMA, 278 (1997), 58–62.

[3] Henrard, D. R., J. F. Phillips, L. R. Muenz, W. A. Blattner, D. Weisner, M. E. Eyster, and J. J. Goedert, *Natural history of HIV-1 cell-free viremia*, JAMA, 274 (1995), 554–558.

[4] O'Brien, T. R., W. A. Blattner, D. Waters, M. E. Eyster, M. W. Hilgartner, A. R. Coher, N. Luban, A. Hatzakis, L. M. Aledort, P. S. Rosenberg, W. J. Milet, B. L. Kroner, and J. J. Goedert, *Serum HIV-1 RNA levels and time to development of AIDS in the multicenter hemophilia cohort study*, JAMA, 276 (1996), 105–110.

[5] Wong, M. T., M. J. Dolan, E. Kozlow, R. Doe, G. P. Melcher, D. S. Burke, R. N. Boswell, and M. Vahey, *Patterns of virus burder and T cell phenotype are established early and are correlated with the rate of disease progression in HIV type 1 infected persons*, J. Infect. Dis., 173 (1996), 877–887.

[6] Piatak, M., M. S. Saag, L. C. Yang, S. J. Clark, J. C. Kappes, K. C. Luk, B. H. Hahn, G. M. Shaw, and J. D. Lifson, *High levels of HIV-1 in plasma during all stages of Infection determined by competitive PCR*, Science, 259 (1993), 1749–1754.

[7] Baltimore, D., *Lessons from people with nonprogressive HIV infection*, New England Journal of Medicine, 332 (1995), 259–260.

[8] Cao, Y., L. Qin, L. Zhang, J. Safrit, D. D. Ho, *Virologic and immunologic characterization of long-term survivors of HIV type 1 infection*, N. Engl. J. Med., 332 (1995), 201–208.

[9] Clumeck, N., H. Taelman, P. Hermans, P. Piot, M. Schoumacher, and S. Dewit, *A cluster of HIV infection among heterosexual people without apparent risk-factors,* N. Engl. J. Med., 321 (1989), 1460-1462.

[10] Holmberg, S. D., C. R. Horsburgh, J. W. Ward, and H. W. Jaffe, *Biologic factors in the sexual transmission of human immunodeficiency virus,* J. Infect. Dis., 160 (1989), 116–122.

[11] Palenicek, J., R. Fox, J. Margolick, H. Farzadegan, D. Hoover, N. Odaka, S. Rubb, H. Armenian, J. Harris, and A. J. Saah, *Longitudinal study of homosexual couples discordant for HIV-1 antibodies in the baltimore MACS study*, J. AIDS, 5 (1992), 1204–1211.

[12] Peterman, T. A., R. L. Stonebrunner, J. R. Allen, H. W. Jaffe, and J. W. Curran, *Risk of human immunodeficiency virus transmission from heterosexual adults with transfusion-associated infections,* JAMA. 256 (1988), 55–58.

[13] Ragni, M. V., L. A. Kingsley, P. Nimorwicz, P. Gupta, and C. R. Rinaldo, *HIV heterosexual transmission in hemophilia couples: Lack of relation to T4 number, clinical diagnosis, or duration of HIV exposure,* J. AIDS, 2 (1989), 557–563.

[14] Lazzarin, A., A. Saracco, M. Musicco, and A. Nicolosi, *Man-to-woman sexual transmission of the human immunodeficiency virus*, Arch. Intern. Med., 151 (1991), 2411–2416.

[15] Royce, R. A., A. Sena, W. Cates, and M. S. Cohen, *Sexual transmission of HIV* N. Eng. J. Med., 336 (1997), 1072–1078.

[16] Biti, R., R. French, J. Young, J. Bennets, G. Stewart, T. Liang, *HIV-1 infection in an individual homozygous for the CCR5 deletion allele*, Nat. Med., 3 (1997), 252.

[17] Huang, Yaoxing, et al., *The role of a mutant CCR5 allele in HIV-1 transmission and disease progression,* Nature Medicine, 2 (1996), 1240–1243.

[18] Longini, I. M., W. S. Clark, M. Haber, and R. Horsburgh, *The stages of HIV infection: Waiting times and infection transmission probabilities*, In: Mathematical Approaches to AIDS Epidemiology (Castillo-Chavez, Levin, and Shoemaker, Eds.), Lecture Notes in Biomathematics Vol. 83, Springer-Verlag, New York (1989), 111–137.

[19] Anderson, R. M., R. M. May, G. F. Medley, and A. Johnson, *A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS,* IMA, J. Math. Med. Biol., 3 (1986), 229–263.

[20] Jacquez, J. A., C. P. Simon, J. Koopman, L. Sattenspiel, and T. Perry, *Modelling and analyzing HIV transmission: The effect of contact patterns,* Math. Biosc., 92 (1988), 119–199.

[21] Jacquez, J. A., J. S. Koopman, C. P. Simon, and I. M. Longini, *Role of the primary infection in epidemics of HIV infection in gay cohorts,* J. AIDS, 7 (1994), 1169–1184.

[22] Lin, Xiaodong, *Qualitative analysis of an HIV transmission model,* Math. Biosc., 104 (1991), 111-134.

[23] Lin, Xiaodong, H. W. Hethcote, and P. Van den Driessche, *An epidemiological model for HIV/AIDS with proportional recruitment,* Math. Biosc., 118 (1993), 181-195.

[24] Thieme, H. R. and C. Castillo-Chavez, *How may infection-age dependent infectivity affect the dynamics of HIV/AIDS?*, SIAM J. Appl. Math., 53 (1993), 1449–1479.

[25] Wiley, J. A., S. J. Herschkorn, and N. S. Padian, *Heterogeneity in the probability of HIV transmission per sexual contact: the case of male-to-female transmission in penile-vaginal intercourse,* Stat. in Med., 8 (1989), 93–102.

[26] Ho, D. D., M. Tarsem, and A. Masud, *Quantitation of human immunodeficiency virus type 1 in the blood of infected persons*, N. Engl. J. Med., 321 (1989), 1621–1625.

[27] Pedersen, C., C. M. Nielsen, B. F. Vestergaard, J. Gerstoft, K. Krogsgaard, and J. O. Nielsen, *Temporal relation of antigenaemia and loss of antibodies to core antigens to development of clinical disease in HIV infection*, Br. Med. J., 295 (1987), 567–569.

[28] Allian, J. P., Y. Laurian, D. A. Paul, and D. Senn, *Serological markers in early stages of immunodeficiency virus infection in haemophiliacs,* Lancet, 2 (1986), 1233-1236.

[29] Sydow, M., H. Gaines, A. Sonnerborg, M. Forsgren, P. O. Pehrson, and O. Strannegard, *Antigen detection in primary HIV infection,* Br. Med. J., 296 (1989), 238–240.

[30] Jacquez, J. A., C. P. Simon, and J. Koopman, *Core groups and the $R_0$s for subgroups in heterogeneous SIS and SI models,* In: Epidemic Models: Their Structure and Relation to Data (Mollison, Ed.), Cambridge Univ. Press, Cambridge (1995), 279–301.

[31] Kemper, J. T., *On the identification of superspreaders for infectious disease,* Math. Biosci., 48 (1980), 111–127.

[32] Hyman, J. M., Jia Li, and A. E. Stanley, *Sensitivity Studies of the Differentiated Infectivity and Staged Progression Models for the Transmission of HIV*, In preparation, available from http://math.lanl.gov/ mac/papers/papers.html in late 1998.

[33] Mellors, J. W., A. Munoz, J. V. Giorgi, J. B. margolick, C. J. Tassoni, P. Gupta, L. A. Kingsley, J. A. Todd, A. J. Saah, R. Detels, J. P. Phair, and C. R. Rinaldo, Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection, Ann. Intern. Med., 126 (1997), 946–954.

[34] Gupta, P., J. Mellors, L. Kingsley, S. Riddler, M. Singh, S. Schreiber, M. Cronin, and C. Rinaldo, *High viral load in semen of HIV-1 infected men at all stages of disease and its reduction by antiretroviral therapy*, 4th Conf. Retro. and Opportun. Infect., abstract no. 726 (1997).

[35] Hart, C., M. Palmore, T. Wright, J. Lennox, T. Evans-Strickfaden, T. Bush C. Schnell, L. Conlet, and T. V. Ellerbrock, *Correlation of cell-free and cell-associated HIV RNA levels in plasma and vaginal sectretions*, 4th Conf. Retro. and Opportun. Infect., abstract no. 25 (1997).

[36] Speck, C. E., R. Coombs, L. Koutsky, J. Zeh, L. Corey, T. Hooton, S. Ross, and J. Krieger, *Rates and determinants of HIV shedding in semen*, Int. Conf. AIDS, abstract no. We. C.334 (1996).

[37] Jurrianns, S., P. L. Vernazza, J. Goudsmit, J. Boogaard, B. Van Gemen, *HIV-1 viral load in semen versus blood plasma*, Int. Conf. AIDS, abstract no. Th.A.4036 (1996).

[38] Dickover, R. E., E. M. Garratty, S. A. Herman, M. S. Sim, S. Plaeger, P. J. Boyer, M. Keller, A. Deveikis, E. R. Stiehm, and Y. J. Bryson, *Identification of levels of maternal HIV-1 RNA associated with risk of perinatal transmission*, JAMA, 275 (1996), 599–605.

[39] Padian, N., L. Marquis, D. P. Francis, R. E. Anderson, G. W. Rutherford, P. M. O'Malley, and W. Winkelstein, *Male-to-female transmission of human immunodeficiency virus*, JAMA, 258 (1987), 788–790.

[40] MMWR, *Update: transmission of HIV infection during and invasive dental procedure – Florida*, 40 (1991), 21–27.

[41] Pezzotti, P., A. N. Phillips, M. Dorrucci, A. C. Lepri, N. Galai, D. Vlahov, and G. Rezza, *Category of exposure to HIV and age in the progression to AIDS*, Br. Med. J., 313 (1996), 583–586.

[42] Latif, A. S., D. A. Katzenstein, M. T. Basset, S. Houston, J. C. Emmanuel, and E. Marowa, *Genital ulcers and transmission of HIV among couples in Zimbabwe*, J. AIDS, 3 (1989), 519–523.

[43] Royce, R. A., W. Winkelstein, and P. Bachetti, *Cigarette smoking and incidence of AIDS*, Abstracts of the Sixth International Conference on AIDS, San Francisco, CA (1990), Abstract Th.C.39.

[44] Henin, Y., L. Mandelbrot, R. Henrion, R. Pradinaud, J. P. Couland, and L. Montagnier, *Virus excretion in the cervicovaginal secretions of pregnant and nonpregnant HIV-infected women*, J. AIDS, 6 (1993), 72–75.

[45] Landesman, S. H. and D. Burns, *Quantifying HIV*, JAMA, 275 (1996), 640–641.

[46] Garred, P., *Susceptibility to HIV infection and progression of AIDS in relation to variant alleles of Mannose-Binding lectin*, Lancet, 349 (1997), 236–240.

[47] Greenblatt, R. M., S. A. Lukehart, F. A. Plummer, T. C. Quinn, C. W. Critchlow, R. L. Ashley, L. J. D'Costa, J. O. Ndinya-Achola, L. Corey, A. R. Ronald, and K. K. Holmes, *Genital ulceration as a risk factor for HIV infection*, J. AIDS, 2 (1988), 47–50.

# Appendixes

## A The Reproductive Number for the DI Model

The asymptotic dynamics of the infection-free equilibrium are determined by the following submatrix in the Jacobian matrix of (2.1) at the infection-free equilibrium

$$J := \begin{pmatrix} p_1 r\beta_1 - (\mu+\nu_1) & p_1 r\beta_2 & \dots & p_1 r\beta_n \\ p_2 r\beta_1 & p_2 r\beta_2 - (\mu+\nu_2) & \dots & p_2 r\beta_n \\ \vdots & \vdots & \ddots & \vdots \\ p_n r\beta_1 & p_n r\beta_2 & \dots & p_n r\beta_n - (\mu+\nu_n) \end{pmatrix}.$$

Since all off-diagonal elements are positive, we now consider matrix $-J$. Using the positive vector $V := \left(\frac{p_1}{(\mu+\nu_1)} \cdots \frac{p_n}{(\mu+\nu_n)}\right)^T$, we have

$$-J \cdot V = \left(1 - r\sum_{j=1}^{n} \frac{p_j\beta_j}{(\mu+\nu_j)}\right) E,$$

where $E = (p_1, p_2, \cdots, p_n)^T$. Then it follows from M-matrix theory that each eigenvalue of $J$ has a negative real part if

$$R_0 = r\sum_{j=1}^{n} \frac{p_j\beta_j}{(\mu+\nu_j)} < 1,$$

which implies the infection-free equilibrium is locally asymptotically stable.

On the other hand, by mathematical induction, it can be shown that the determinant of $J$ equals

$$(-1)^{n+1}\prod_{i=1}^{n} \frac{\mu+\nu_i}{\beta_i}\left(R_0 - 1\right).$$

Then, if $R_0 > 1$, $J$ has eigenvalues with positive real parts, which implies the instability of the infection-free equilibrium.

## B Stability of the Endemic Equilibrium for the DI Model

For the DI model, we establish the local stability of the endemic equilibrium by showing that all eigenvalues of the Jacobian matrix at the endemic equilibrium have negative real parts.

The Jacobian matrix at the endemic equilibrium has the form

$$J = \begin{pmatrix} -\mu - \lambda\left(1 - \frac{S}{N}\right) & -h_1 & -h_2 & \dots & -h_n \\ p_1\lambda\left(1 - \frac{S}{N}\right) & -(\mu+\nu_1) + p_1 h_1 & p_1 h_2 & \dots & p_1 h_n \\ p_2\lambda\left(1 - \frac{S}{N}\right) & p_2 h_1 & -(\mu+\nu_2) + p_2 h_2 & \dots & p_2 h_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_n\lambda\left(1 - \frac{S}{N}\right) & p_n h_1 & p_n h_2 & \dots & -(\mu+\nu_n) + p_n h_n \end{pmatrix},$$

where $h_i := \dfrac{S}{N}(r\beta_i - \lambda)$.

By the similarity matrix

$$M := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ p_1 & 1 & 0 & \dots & 0 \\ p_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_n & 0 & 0 & \dots & 1 \end{pmatrix},$$

$J$ is transformed into the following matrix:

$$B := \begin{pmatrix} -\mu - \lambda + \frac{r}{R_0}\sum\limits_i p_i\beta_i & -h_1 & -h_2 & \dots & -h_n \\ p_1\nu_1 & -(\mu+\nu_1) & 0 & \dots & 0 \\ p_2\nu_2 & 0 & -(\mu+\nu_2) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_n\nu_n & 0 & 0 & \dots & -(\mu+\nu_n) \end{pmatrix}.$$

Hence, matrix $J$ has all eigenvalues with negative real parts if and only if $B$ has all eigenvalues with negative real parts.

Let $\rho$ be an eigenvalue of $B$ with the corresponding eigenvector $X = (x_0, x_1, \cdots, x_n)^T$. Then

$$p_j\nu_j x_0 = (\mu + \nu_j + \rho)\, x_j, \qquad j = 1, \cdots, n. \tag{B.1}$$

Solving (B.1) for $x_j$, $j = 1, \cdots, n$, in terms of $x_0$ and then substituting them into the first row of $B\mathbf{X} = \rho\mathbf{X}$, we arrive at

$$-\mu - \lambda + \frac{r}{R_0}\sum_{j=1}^{n} p_j\beta_j - \frac{1}{R_0}\sum_{j=1}^{n}\frac{p_j\nu_j\,(r\beta_j - \lambda)}{(\mu + \nu_j + \rho)} = \rho.$$

Let $\rho = u + iv$. Then $u$ and $v$ satisfy the following equations:

$$u + \mu + \lambda - \frac{r}{R_0}\sum_{j=1}^{n} p_j\beta_j + \frac{1}{R_0}\sum_{j=1}^{n} \frac{p_j\nu_j\,(r\beta_j - \lambda)\,(u + \mu + \nu_j)}{(u + \mu + \nu_j)^2 + v^2} = 0, \tag{B.2a}$$

$$1 - \frac{1}{R_0}\sum_{j=1}^{n} \frac{p_j\nu_j\,(r\beta_j - \lambda)}{(u + \mu + \nu_j)^2 + v^2} = 0. \tag{B.2b}$$

Now we have the following two lemmas.

**Lemma B.1.** *Any eigenvalue of $B$ with a nonnegative real part must be a real number.*

*Proof.* Assume $u \geq 0$ and $v > 0$. Then it follows from (B.2b) that

$$\begin{aligned} R_0 + \lambda \sum_{j=1}^{n} \frac{p_j\nu_j}{(u + \mu + \nu_j)^2 + v^2} =& r\sum_{j=1}^{n} \frac{p_j\nu_j\beta_j}{(u + \mu + \nu_j)^2 + v^2} \\ &< r\sum_{j=1}^{n} \frac{p_j(\mu + \nu_j)\beta_j}{(u + \mu + \nu_j)^2} \leq r\sum_{j=1}^{n} \frac{p_j(\mu + \nu_j)\beta_j}{(\mu + \nu_j)^2} = R_0, \end{aligned}$$

which is impossible. Hence, if $u \geq 0$, $v$ must be nonpositive. However, since $u - iv$ is also an eigenvalue of $B$, $v$ cannot be negative. Hence, $v$ must be zero, and $\rho$ must be a real number. □

**Lemma B.2.** *Matrix $B$ has no nonnegative real eigenvalue.*

*Proof.* Matrix $B$ has a positive real eigenvalue $\rho = u$ if and only if (B.2a) has a positive solution $u$ while $v = 0$.

Define

$$F(u) := u + \mu + \lambda - \frac{r}{R_0}\sum_{j=1}^{n} p_j\beta_j + \frac{1}{R_0}\sum_{j=1}^{n} \frac{p_j\nu_j\,(r\beta_j - \lambda)}{u + \mu + \nu_j}.$$

Then

$$F(0) = \mu + \lambda - \frac{r}{R_0}\sum_{j=1}^{n} p_j\beta_j + \frac{1}{R_0}\sum_{j=1}^{n} \frac{p_j\nu_j\,(r\beta_j - \lambda)}{\mu + \nu_j} \geq \lambda\left(1 - \frac{1}{R_0}\right) > 0,$$

if $R_0 > 1$.

On the other hand,

$$\begin{aligned} F'(u) =& 1 - \frac{1}{R_0}\sum_{j=1}^{n} \frac{p_j\nu_j\left(r\beta_j - \lambda\right)}{\left(u+\mu+\nu_j\right)^2} \geq 1 - \frac{r}{R_0}\sum_{j=1}^{n} \frac{p_j\nu_j\beta_j}{\left(u+\mu+\nu_j\right)^2} \\ &\geq 1 - \frac{r}{R_0}\sum_{j=1}^{n} \frac{p_j(\mu+\nu_j)\beta_j}{\left(u+\mu+\nu_j\right)^2} \geq 1 - \frac{r}{R_0}\sum_{j=1}^{n} \frac{p_j\beta_j}{\mu+\nu_j} = 0, \end{aligned}$$

for all $u \geq 0$. Hence, $F(u) > 0$ for all $u \geq 0$. That is, matrix $B$ has no nonnegative real eigenvalue. □

Based on Lemmas B.1 and B.2, all eigenvalues of $B$ must have negative real parts. Then the local asymptotic stability of the endemic equilibrium of (2.1) follows.

# C The Endemic Equilibrium for the SP Model

Defining $\sigma_i = \gamma_i + \mu$, it follows from the equations for $I_i$, $i = 2, \cdots, n$, in (3.1) that

$$I_{i-1} = \frac{\sigma_i}{\gamma_{i-1}} I_i.$$

Define

$$\Delta_i := \frac{\prod\limits_{j=i+1}^{n} \sigma_j}{\prod\limits_{j=i}^{n-1} \gamma_j}.$$

Then

$$I_i = \Delta_i I_n, \qquad i = 1, \cdots, n-1. \tag{C.1}$$

Since $\lambda S = \sigma_1 I_1$, $Sr\sum\limits_{k=1}^{n}\beta_k I_k = S\sigma_1 I_1 + \sigma_1 I_1 \sum\limits_{k=1}^{n} I_k$. Then $Sr\sum\limits_{k=1}^{n}\beta_k\Delta_k = S\sigma_1\Delta_1 + \sigma_1\Delta_1\sum\limits_{k=1}^{n}\Delta_k I_n$, and hence

$$S = \frac{\sigma_1\Delta_1\sum\limits_{k=1}^{n}\Delta_k}{r\sum\limits_{k=1}^{n}\beta_k\Delta_k - \sigma_1\Delta_1} I_n = \frac{\sum\limits_{k=1}^{n}\Delta_k}{\frac{r}{\Delta_1\sigma_1}\sum\limits_{k=1}^{n}\beta_k\Delta_k - 1} I_n.$$

It can be shown that

$$R_0^S = \frac{r \sum_{k=1}^{n} \beta_k \Delta_k}{\sigma_1 \Delta_1}.$$

Then,

$$S = \frac{\sum_{k=1}^{n} \Delta_k}{R_0^S - 1} I_n, \tag{C.2}$$

which is positive if and only if $R_0^S > 1$.

On the other hand, from $\mu(S^0 - S) = \sigma_1 I_1 = \sigma_1 \Delta_1 I_n$, it follows that

$$\mu S^0 = \left( \frac{\mu \sum_{k=1}^{n} \Delta_k}{R_0^S - 1} + \sigma_1 \Delta_1 \right) I_n.$$

Thus

$$I_n = \frac{\mu S^0}{\frac{\mu \sum_{k=1}^{n} \Delta_k}{R_0^S - 1} + \sigma_1 \Delta_1}. \tag{C.3}$$

Substituting (C.3) into (C.2) and (C.1), respectively, we can solve for the endemic equilibrium explicitly as in (3.8).