

HIV Epidemics Driven by Late Disease Stage Transmission

Brandy L. Rapatski, PhD,[†] Frederick Suppe, PhD,^{*§} and James A. Yorke, PhD^{*†‡}

Summary: How infectious a person is when infected with HIV depends on what stage of the disease the person is in. We use 3 stages, which we call primary, asymptomatic, and symptomatic. It is important to have a systematic method for computing all 3 infectivities so that the measurements are comparable. Using robust modeling, we provide high-resolution estimates of semen infectivity by HIV disease stage. We find that the infectivity of the symptomatic stage is far higher, hence more potent, than the values that prior studies have used when modeling HIV transmission dynamics. The stage infectivity rates for semen are 0.024, 0.002, and 0.299 for the primary, asymptomatic, and symptomatic stages, respectively. Implications of our infectivity estimates and modeling for understanding heterosexual epidemics such as that in sub-Saharan African are explored.

Key Words: HIV, infectiousness, transmission dynamics, mathematic model, effective contact rate

(*J Acquir Immune Defic Syndr* 2005;38:241–253)

Our goal is to determine how the infectiousness of semen of HIV-infected men varies by stage of disease. The actual infectiousness varies from person to person; thus, we compute an average. It also depends on the type of sexual act and is higher for some types than for others. We find that susceptible gay men involved in unprotected receptive anal intercourse (RAI) are 12.5 times more likely to become infected when the partner is an infected symptomatic man than when the partner is in the primary stage; that is, symptomatic men are 12.5 times as infectious as men in the primary stage. Furthermore, men in the asymptomatic stage are 149.5 times less infectious than symptomatic men. We expect this general pattern of infectivity to hold for all sex acts of infected men.

We focus on measuring the infectivity of gay men. By the “infectivity” of a person, we mean the fraction of his susceptible “contacts” that he infects. For HIV, we interpret a “contact” to be the activity during which most infected gay men became infected.^{1,2} Infectivity varies as the disease prog-

resses in the individual, and, of course, one cannot conduct experiments to determine infectivity. In this article, we analyze the San Francisco City Clinics Cohort (SFCCC) data and figure out how many contacts there are at each point in time between susceptible and infected men and what stages of infection those men are in. This is not a straightforward calculation. We take into account the great variability between the contact rates of the men.

SAN FRANCISCO CITY CLINICS COHORT DATA

The only high-resolution data set documenting the onset of HIV in a population is the SFCCC study, which is based on blood samples from an earlier hepatitis B vaccine clinical trial that took place during the period in which HIV exploded throughout the San Francisco gay population (Table 1).^{3–5} That study, involving approximately 10% of the San Francisco gay population, involved 6875 men and took blood samples and behavioral data. After HIV was identified, stored blood samples were thawed and tested for the presence of HIV antibodies. This enabled documentation of the growth of HIV throughout the population (Fig. 1) and relation of that growth to behavioral data. We stress that with respect to HIV incidence, that data set is a biologic one not dependent on medical diagnosis or infection self-report.⁴

VARIABLE INFECTIVITY

It is widely understood that epidemiologic modeling of HIV transmission must use at least 3 stages (primary infection, asymptomatic, and symptomatic, including AIDS) with different infectivities for each stage (Fig. 2). First comes a period of primary infection (lasting part of a year). Our “primary infectious stage” is defined as the time soon after initial infection when infectiousness first rises and then drops. Seroconversion typically occurs well before the end of our primary stage. One then enters an asymptomatic period (averaging 7–8 years without treatment) during which infectiousness is low, followed by a symptomatic stage (averaging 3 years until death without treatment) during which infectiousness rises again. The symptomatic stage begins while individuals are relatively healthy and active, although it also includes the more severe AIDS phase. These average times are based on SFCCC data.⁶

Viral levels also vary greatly between these 3 stages. During the period of primary infection, viral levels are typically high. The viral levels become low as one enters the asymptomatic period, followed by a symptomatic/AIDS stage during which the viral loads are extremely high (Fig. 3).^{7–9}

If one assumes, plausibly, that HIV infectivity correlates with semen viral levels,^{10–13} one would expect that HIV

Received for publication September 16, 2003; accepted September 14, 2004. From the *Institute for Physical Sciences and Technology, University of Maryland, College Park, MD; †Department of Mathematics, University of Maryland, College Park, MD; ‡Department of Physics, University of Maryland, College Park, MD; and §Classical and Modern Languages and Literature, Texas Tech University, Lubbock, TX.

Support provided by the National Science Foundation under grant 0104087 and Texas Tech University.

Reprints: Brandy L. Rapatski, Department of Mathematics, University of Maryland, College Park, MD 20742 (e-mail: brandyrapatski@hotmail.com).

Copyright © 2005 by Lippincott Williams & Wilkins

TABLE 1. SFCCC HIV Onset Data

Column 1	2	3	4	5
Year	% Population Susceptible at Year End	% Population Infected During Year	% Population Symptomatic at Year End	Estimated % Population Symptomatic at Mid-year
1978	95.5			
1979	85.9	9.6	0	
1980	73.8	12.1	0.1	0.05
1981	68.6	5.2	0.55	0.32
1982	51.5	17.1	1.72	1.13
1983	39.3	12.2	3.56	2.64
1984	30.8	8.5	5.85	4.7

Susceptible and infected fractions (columns 2 and 3) for each year (column 1) are from the published source.^{3,4} The percentage of the population symptomatic at year end (column 4) is derived from the optimized solution of the stage/substage model of HIV transmission given in Figure 11. The data show reasonable agreement with 1987 estimate,¹⁴ doubled to correct for the 1993 redefinition of AIDS.⁴⁵ The estimated percentage of the population symptomatic at mid-year (column 5) is the average of this and the previous year's year-end symptomatic fractions.

transmission is more infectious during the primary infection and symptomatic stages than during the asymptomatic stage and is even more infectious during the symptomatic stage. Our model does not use information about viral loads but obtains results showing that infectivity follows a pattern similar to that of the viral loads shown in Figure 3.

The fact that the primary stage played a significant role in the original gay HIV epidemic is widely understood as a result of epidemiologic modeling by Ahlgren et al,⁶ Hethcote and Van Ark,¹⁴ Longini et al,¹⁵ Jacquez et al,¹⁶ and others. None

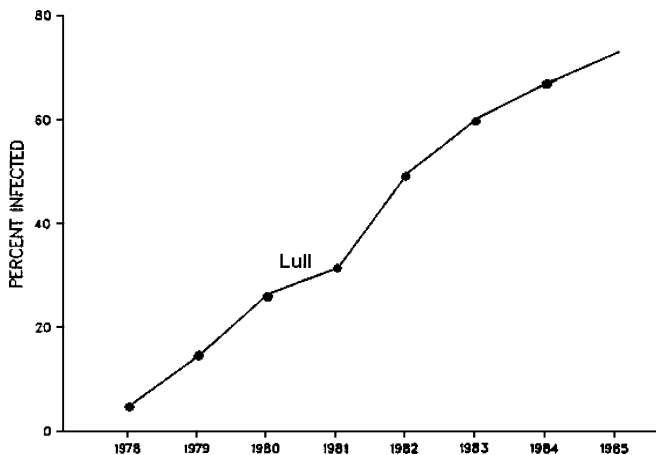


FIGURE 1. San Francisco City Clinics Cohort (SFCCC) epidemic. Growth of HIV infection through the SFCCC.⁴ Note the “lull” from 1980 to 1981, wherein the epidemic slows down before explosively restarting the next year. This phenomenon in the data has not been addressed in the epidemiologic literature modeling the SFCCC epidemic data. (It may be a statistical artifact.) Our modeling shows that this lull can be reproduced only in variable infectivity models in which late-stage (symptomatic) transmission dominates the epidemic once approximately 30% of highly promiscuous populations are infected.

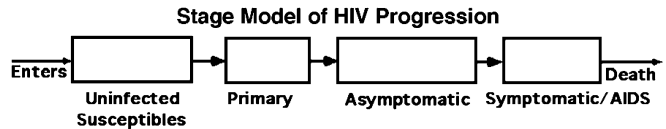


FIGURE 2. HIV stages⁴⁴ and infectiousness. A susceptible becomes newly infected and progresses from a primary infection stage through an asymptomatic stage to a symptomatic stage, followed by death. Infectiousness varies greatly from stage to stage.

of these prior studies attributes a significant role to late-stage symptomatic transmission in shaping the epidemic. A frequently referenced article¹⁶ states that “a review of the data on infectivity per contact for transmission of the HIV suggests that the infectivity may be on the order of 0.1 to 0.3 per anal intercourse in the period of the initial infection, 10^{-4} to 10^{-3} in the long asymptomatic period, and 10^{-3} to 10^{-2} in the period leading to AIDS.” Jacquez et al¹⁶ obtained the primary and asymptomatic stage infectivities by looking at the initial growth of the epidemic and determined what infectivities would be necessary to create such exponential growth. In contrast to those results, which were carefully obtained via detailed models, they say, without giving any details, that the symptomatic stage infectivity was estimated using previous partner studies¹⁷⁻¹⁹ and, in most cases, heterosexual partner studies. It is unclear how those stage estimates were obtained, because those studies do not differentiate infectivity per stage but had an “average” infectivity for the course of the disease. We remark that in the study by Peterman et al (Fig. 1, page 56),²⁰ one can see that there is a negative correlation between the number of contacts a couple has and the probability that the disease will be transmitted. For couples with more contacts, it was reported less likely for the susceptible partner to become infected. Ahlgren et al⁶ reported that they were unable to obtain reliable estimates of symptomatic stage infectivity using data on the reported incidence of AIDS infections. In the early epidemic, those data are unreliable because of underreporting, misdiagnosis, and changing definitions of AIDS in estimating symptomatic stage infectivities. Our approach relies neither on those prior partner study estimates nor on problematic reported AIDS incidence data.

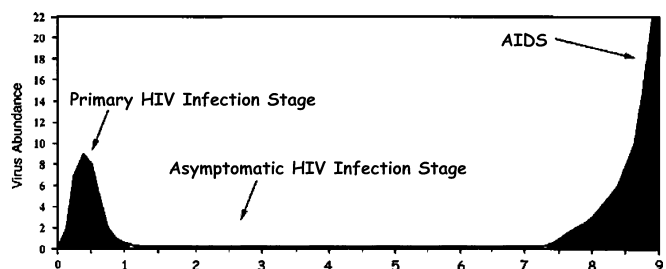


FIGURE 3. Viral loads. Schematic representation of the variation in viral loads over the course of a typical untreated individual's HIV infection (Adapted from Anderson RM. The spread of HIV and sexual mixing patterns. In: Mann J, Tarantola D, eds. AIDS in the World II: Global Dimensions, Social Roots, and Responses. The Global AIDS Policy Coalition. New York: Oxford University Press; 1996:71-86.)

ACTIVITY LEVELS

The SFCCC reports annual numbers of partners for 6 different activity levels.^{6,9} The most active half have more than 10 times as many contacts as the lower half. By 1982, almost half of the population was infected, presumably primarily the most active half. The most active 10% (the “core”) are responsible for nearly half of all sexual contacts.²¹

Published survey data (Table 2, column 3) give the distribution of the number of partners for 6 activity levels but do not give RAI group averages. Nevertheless, the data indicate that for the overall population, 57% of the contacts involved RAI by at least 1 of the 2 partners^{6,22}; thus, we estimate RAI activity for each group at 57% of total activity.* These estimates are shown in column 3 of Table 2.

Most prior (homosexual) studies^{6,9,24} did not differentiate activity groups or else had only 2 groups, a highly active core group consisting of 5% to 10% of the population and a less active group. Models with only 1 or 2 activity groups give results that differ significantly from the 6-group model.

Available evidence suggests that unprotected sexual activity did not decrease significantly until 1985.²⁵ For this reason, we model the epidemic through 1984 with no behavioral change. We also assume that each individual’s sexual activity level is the same for all 3 stages of the infection. We later discuss the effect of modifying these 2 modeling assumptions.

MIXING PATTERNS

We model interactions obtained from published SFCCC survey data among 6 activity groups as if all sexual contacts were casual and promiscuous, such as was typical in gay bathhouses, where contacts are fairly indiscriminate and casual. Our bathhouse assumption is that gay contacts resemble bathhouse patterns. More specifically, published SFCCC data give average contact rates for 6 different activity levels. We assume that these specify the average frequency with which a person goes to the bathhouse; however, once inside, the mixing pattern is random. Thus, the bathhouse assumption addresses the frequency of contacts and the mixing pattern when contacts occur. Because the core has nearly half of all contacts, nearly half of the men in the bathhouse at any given time are in the highly active promiscuous core. Figure 4 shows that there are fewer susceptibles in the bathhouse than in the general population, which an essential fact in understanding the epidemic.

A key to understanding the dynamics of an HIV epidemic is estimating from population totals how different activity level groups of uninfected people become infected (see Fig. 4). Our modeling reveals that the core group rapidly becomes infected early in the epidemic.²⁶ The infection spreads through other groups more slowly. Core members are responsible for 48.5% of all sexual contacts (see Table 2). Once most of the core group is infected, new infections predominantly are of

TABLE 2. Group RAI Contact Behavior

Column 1	2	3	4
Group j	Size F _j	Contacts per Year C _j	Group Fraction of All Contacts
1 (core)	10%	231	0.48
2	15%	81	0.25
3	25%	33	0.17
4	25%	15	0.078
5	15%	3	0.009
6	10%	0	0.000
Total population	100%	48	

Group contact data (columns 1–3) are from the study by Hethcote and Van Ark.¹⁴ Group fraction of all contacts is the total number of annual RAI contacts from the group (column 2 × column 3) divided by the population average of 48.

C_j indicates the number of contacts a person in group j has per year; F_j, fraction of population in group j.

group 2 men until they too are mostly infected. Most new infections are then of group 3 men, and so on (see Fig. 4).

INFECTIVITY ESTIMATES

The only transmission vector for homosexual transmission of HIV that has been shown to be epidemiologically significant is via RAI.²⁷ It is standard to interpret HIV infectivities for gay men as RAI infectivities. Early studies that did not consider difference in contact rates and did not estimate different infectivities for the 3 stages came up with a 1% infectivity for RAI. It is not unreasonable to construe these estimates as the average infectivity of an individual over

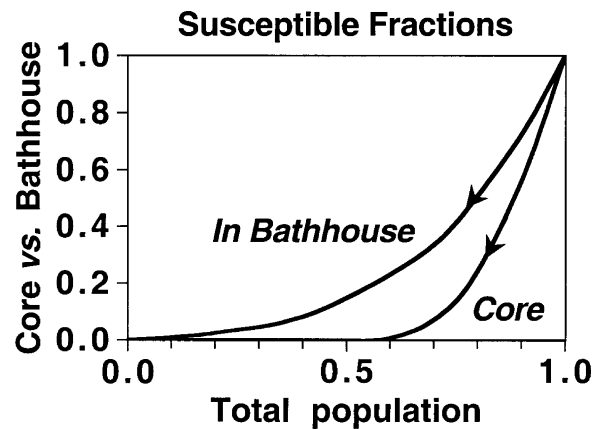


FIGURE 4. Susceptibles in more active groups are soon depleted. An essential finding of the model solution is that there are fewer susceptibles in the bathhouse than in the general population and that even fewer members of the most active groups present in the bathhouse are susceptible. For example, if 60% of the population is susceptible (as was the case in mid-1982), we can conclude that 23% of the men in the bathhouse are susceptible and virtually none of the core members present are susceptible. The method used in this figure represents a reinterpretation of the results presented in Figure 11.

*If the RAI is for only 1 partner, we are overestimating the number of contacts by a factor of 2, resulting in a 50% underestimation of stage infectivities, but not affecting the ratios of stage infectivities. Thus, we divide the population into 6 groups based on this estimated average RAI activity (contacts). The average number of RAI partners per year is 48.^{6,23}

the course of his HIV infection (ie, the average of the stage infectivities weighted by stage durations).

BASIC MODEL

Our basic modeling principle is that the constants in our model should be based on epidemiologic data. We avoid a priori assumptions not supported directly or indirectly by data.

Obtaining “Best-Fit” Infectivities for the 3 Stages

Estimates of average infectiousness for each of the 3 stages are obtained using the Figure 5 model based on Figure 2. The model can be run for any choice of the 3 infectivities, and an epidemic is produced. In particular, the cumulative fraction infected is reported for each of the 7 years from 1978 to 1984. For each of the 7 years, we compute the square of the difference between this model epidemic and the actual SFCCC epidemic. Let RMS denote the square root of the average of those 7 numbers. We use a minimization technique to select the choice of infectivities for which the RMS is minimized.^{4,10,28} The minimum is obtained for stage infectivity rates 0.024, 0.002, and 0.299, respectively (Fig. 6), with an RMS of 0.016. We call these infectivity estimates the “best-fit infectivities.” The model solution displayed in Figure 7 reproduces the “cumulative SFCCC epidemic” using these best-fit infectivities.

Epidemic for the Best-Fit Infectivities

Figure 7 shows that before 1980, approximately 98% of the infections were caused by primary stage men who had only been infected for a few months.† By 1981, this fast transmission wave ended, when most of the highly active men were infected and were in the asymptomatic stage. After 1981, these men begin entering the symptomatic stage and caused most of the new cases. This second wave was a slow transmission wave in which the infectors were mostly symptomatic men who had been infected for years and were now highly infectious. The lull from 1980 to 1981,³ shown in Figure 1, if not an artifact, is consistent with low infectivity for asymptomatic men. Symptomatic men are more than 12 times as infectious as men in the primary stage for perhaps 4 times as long, making the gay slow transmission wave extremely lethal. Figures 7 and 8 reveal the epidemic's structure, which reflects the combined effects of the displayed fast transmission wave before 1980, followed by the displayed slow transmission wave.

†We can understand much of the early dynamics without finding a solution to the model. Of course, any active man could have become infected and then infected others, but it is instructive to focus on the core group and only the infections the members of the core group caused when in the primary stage. Those men had 231 partners per year or approximately 115 in the primary stage (0.5 year). Forty-eight percent of those contacts ($n = 55$) are with men in the core group. The primary stage infectivity is 2.4%, resulting in 1.3 infections. Some of these infections are in the first quarter of the year, and some are in the second quarter; the average time to infection is 0.25 year. Hence, each quarter of a year, the number of infected men grows by a factor of 1.3. The result of 4 such steps (1 year's worth) is a growth with a factor of 3. We see then that the core primary stage men in San Francisco were able to drive the first (fast) wave of the epidemic.

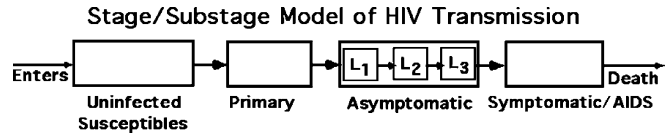


FIGURE 5. Schematic representation of model used for optimized fits to San Francisco City Clinics Cohort epidemic. Three substages have been added to the asymptomatic stage in Figure 2. These substages adjust the standard deviation of the mean time spent in that relatively long stage. Choice of the 3 substages was determined by optimization studies.

IMPLICATIONS OF INFECTIVITY ESTIMATES

Infectiousness of Semen

Our goal has been to determine, using available data, how HIV transmission infectivity varies as the infection progresses within a typical individual not receiving medical treatment. We conclude that symptomatic stage individuals are approximately 12.5 (ie, $0.299/0.024$) times more infectious (per contact) than primary stage men and approximately 149.5 (ie, $0.299/0.002$) times more infectious than asymptomatic stage men. We interpret the infectivities more loosely as measures of the infectiousness of semen in the 3 stages. Doing so enables us to apply our infectivity estimates to heterosexual transmission. Specifically, we would expect rather similar ratios for the stage probabilities of men infecting female contacts. The actual stage infectivities for vaginal intercourse might be higher or lower depending on the type of contact, but the effective stage ratios would be similar. We have no way to measure the corresponding infectiousness of women. In modeling heterosexual populations, we assume that these effective contact ratios are similar for men and women and apply equally across stages.‡

Effective Contact Rates for the Stages

We now take into account the duration of the stage. If all the partners of a man were susceptible, for each contact per year, the number of men we would expect him to infect is the product of infectivity and duration. It is 0.012 men (0.024×0.5) for primary stage men, 0.014 (0.002×7) for asymptomatic stage men, and 0.897 (0.299×3) for symptomatic stage men. We call these the “effective contact rates for the stages.” These numbers measure the relative danger of the 3 stages to all susceptible partners. The effective contact rate for the symptomatic stage is 75 times that of the primary stage ($0.897/0.012$) and approximately 64 ($0.897/0.014$) times that of the asymptomatic stage. Hence, if most of a man's partners are susceptible, during the symptomatic stage, he is likely to infect 75 times as many partners as when he is in the primary stage and 64 times as many as in the latent stage. That does not mean that the primary period is unimportant, because in San Francisco, primary contacts were the main method of

‡Studies of heterosexual populations show that untreated infected individual's viral loads follow a pattern of moderate, then low, and then high as a person progresses through the disease.²⁹⁻³² This corresponds to the pattern of our infectivity estimates.

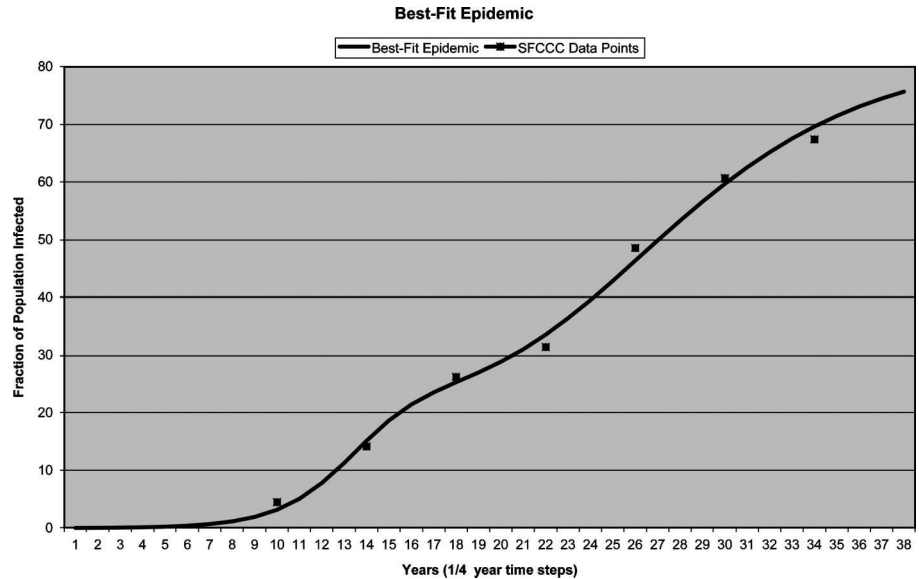


FIGURE 6. Best-fit epidemic. The best-fit epidemic is produced by running our model with the best-fit infectivities (0.024, 0.002, and 0.299). This produces a square root of the average error of 0.016.

transmission up through 1980. Instead, it demonstrates how truly dangerous the symptomatic stage is.

If an average infected man contacts N susceptibles per year, over the course of his infection, he will contact 10.5 because the average duration of infection is 10.5 years. Furthermore, he will infect $0.923N$ men over the course of his infection, where 0.923 is the sum of the effective contact rates

stated previously ($0.012 + 0.014 + 0.897$). We calculate the fraction of the people infected in each stage as follows:

$$P_1 = 0.012/0.923 = 0.013$$

$$P_2 = 0.014/0.923 = 0.015$$

$$P_3 = 0.897/0.923 = 0.972$$

Thus, more than 97% of the infections that a man causes are transmitted when he is in stage 3 (note that $P_1 + P_2 + P_3 = 1.0$). Recall that this is assuming almost all partners are susceptibles, but it also holds when the fraction of partners that are susceptible remains constant, such as when the epidemic is in equilibrium in the population. The often repeated assertion that the primary stage plays the most significant role in driving

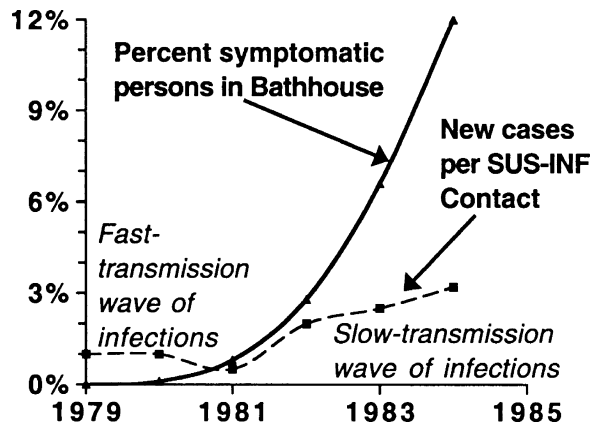


FIGURE 7. The solid curve shows the increase in the number of symptomatic men in the bathhouse where transmission takes place. Symptomatics do not become a significant presence in the bathhouse until 1981, when 31% of the population had been infected during the fast wave. Average "infectiousness" (dashed line) is the percentage of susceptibles infected by a contact with an infected man in the bathhouse (labeled as "new cases per SUSceptible-INFected contact"). Early in the epidemic, when most transmissions are from men in the primary stage, infectiousness is 1%. Later in the epidemic, during the slow transmission wave, infectiousness triples as men in the symptomatic stage begin transmitting HIV. Because they are a small fraction of the infected men in the bathhouse, their average infectiousness must be at least 17% to 28%. The results presented in this figure are based on model equations.

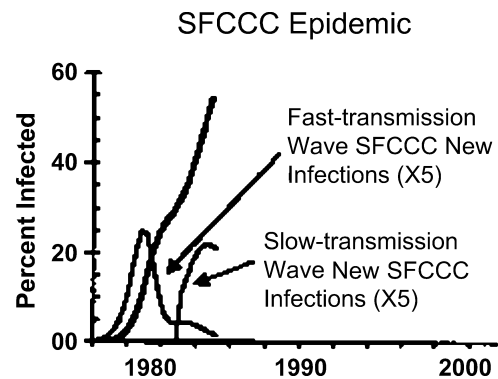


FIGURE 8. Two-wave gay epidemic. Primary stage per-contact infectivity is 2.4%, asymptomatic infectivity is 0.2%, and symptomatic infectivity is 30%. The model has solutions that closely approximate the San Francisco City Clinics Cohort incidence data only when the infectivities are close to these infectivity levels. Men in the primary stage drive the fast transmission wave, and men in the symptomatic stage dominate the slow transmission wave. The results presented in this figure are based on model equations.

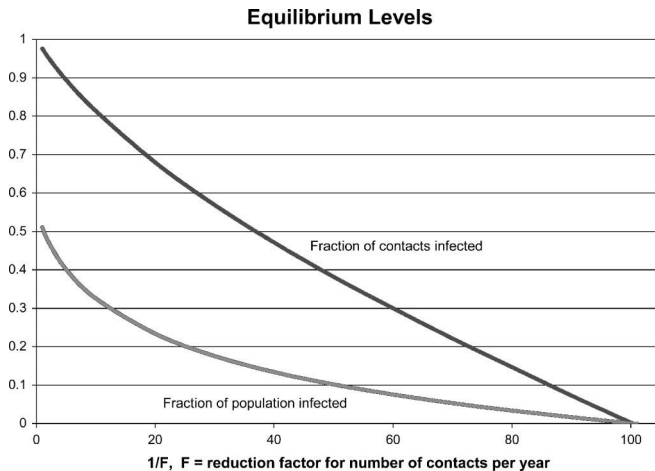


FIGURE 9. Equilibrium levels. At any average annual contact rate over time, an epidemic stabilizes at some fraction of the population infected. What this equilibrium level is depends on the average annual number of contacts. The graph displays equilibrium levels ranging from the San Francisco City Clinics Cohort (SFCCC) data ($F = 1$) to contact rates $1/100$ of the SFCCC data ($F = 100$). Whenever $F < 101$, the equilibrium results in some portion of the population infected; thus, $R_0 = 101$. The SFCCC peak is 70.2%, whereas the equilibrium level is 50%.

an HIV epidemic holds only when the epidemic is growing rapidly, such as it did in San Francisco before 1981.

Mean Transmission Time

When a person is infected, we refer to the “transmission time” as the length of time the infector was infected at that point in time. For example, if someone is infected by a person in his/her primary stage, the transmission time would be approximately 0.25 year on average, whereas if the infector was in stage 2 or 3, the transmission time would be approximately 4.0 or 9.0 years, respectively, on average, representing the time from initial infection to the middle of the stage.

When an epidemic is at equilibrium, the mean transmission time is computed by weighting these 3 times by the fraction of people infected in that stage:

$$(P_1 * 0.25 + P_2 * 4.0 + P_3 * 9.0) = 8.81 \text{ years}$$

If the epidemic is not at equilibrium but is growing, then there will be two possible patterns. In the first, there are relatively more people who were recently infected than infected longer ago. In a growing epidemic, when a susceptible finally meets an infected person, that person is, for example, more likely to be in the primary stage, as happened in San Francisco before 1981 (see Fig. 9). In the second pattern, the majority of infected people come from those in the third stage. This was the case in San Francisco after 1981. A more precise calculation would take into account the exponential rate at which the epidemic is growing.

Sub-Saharan Africa

Estimating mean transmission times allows us to extend our findings to sub-Saharan African and other heterosexual pop-

ulations. It is not known when the first case of HIV occurred in Africa, but it is believed to have been around 1950 or perhaps earlier.³³ From 1950 to 1990, less than 1% of the population was infected as the number of cases went from 1 to approximately 1,000,000, producing a slow transmission wave.³⁴⁻³⁶ We estimate a mean transmission time of approximately $T = 7.44$ years based on our values of infectivities, where an epidemic grows by a factor of 1,000,000 in 40 years. We now can estimate how difficult it would be to stop the epidemic: R_0 is defined to be the average number of secondary cases caused by an average infected individual at the beginning of the epidemic when almost all are susceptible. Stopping the epidemic quickly requires interventions that result in driving R_0 well below 1. Let the number of generations, G , be $40/T \sim 5.4$. If the epidemic grows by a factor of R_0 for each of G generations, the compound growth is R_0^G , which should equal approximately 1,000,000. Our estimate is that $T = 7.44$ yields $R_0 = 13$. This calculation depends on the infectivities and duration of the stages and not on other aspects of our SFCCC model.

If an epidemiologist assumes the epidemic started, for example, in 1930, the argument changes, yielding $R_0 = 4.3$, which still remains well above 1. In either case, major sociologic changes are necessary to drive R_0 below 1. If the epidemiologist was not aware of the importance of the symptomatic stage and assumed that the epidemic was driven by the primary stage, the mean transmission time would be short. For the sake of argument, we take it to be 1 year. There are then 40 generations and R_0 would be 1.4. This incorrect assumption would lead to an underestimation of the severity of the epidemic. Only relatively small changes would be necessary to bring R_0 below 1, stopping the growth of the epidemic.

For San Francisco, we note that a reduction of the effective contact rate by a factor of 101 would have been necessary to prevent the epidemic from growing early on (Fig. 9). In other words, R_0 was 101 for the SFCCC study.

METHODOLOGIC DISCUSSION OF MODELING ASSUMPTIONS

We want to know whether our results are artifacts of our modeling assumptions. In this section, we discuss our preferred use of 6 activity groups in our modeling rather than the 1 or 2 activity groups used by other researchers, the tacit assumption of persisting levels of sexual activity, random mixing on the bathhouse assumption, and the use of deterministic rather than stochastic models.

Six-Group Model Assumption

Earlier variable infectivity models^{6,14-16,24,37,38} used 2 activity levels at most instead of our 6 activity levels. Core versus noncore models capture the importance of primary stage infectivity in initiating the gay epidemic but underestimate and thereby miss the critical role of symptomatic stage infectivity in sustaining and intensifying the gay epidemic after 1980 and in initiating and sustaining slow transmission epidemics, such as that in sub-Saharan Africa. We compared our 6-group assumption by making new optimization fits to the SFCCC incidence curves using 2-group and 1-group models.

TABLE 3. Comparison of Our 6-Group 2-Group, and 1-Group Models

	Average Contacts per Year	Best-Fit Stage Infectivity Estimates			RMS Error of Fit to SFCCC Data Points
		Primary	Asymptomatic	Symptomatic	
Six-group model	See Table 2	0.024	0.002	0.299	0.016
Two-group model A	231: 10% of population 27: 90% of population	0.024	0.005	0.091	0.025
Two-group model B	231: 10% of population 53: 90% of population	0.028	0	0.089	0.037
One-group model	48	0.042	0	0.088	0.035

Note: Error $\sqrt{\frac{1}{7} \sum_{1978-1984} [\text{infected}_{\text{model}} - \text{infected}_{\text{actual}}]^2}$

It follows from the equations for each of these models that if all the contacts per year are changed by a fixed factor, (eg, doubled), all the infectivities are reduced by that factor, (ie, cut in half). The 1-group model produces epidemics that always severely disagree with the San Francisco data and cannot explain slow epidemics such as that in Africa. The 2-group and 1-group models underestimate symptomatic stage and thus cannot explain slow epidemics such as that in Africa.

We use the 6-group model because it maximally uses the available contact data. What happens if we use simplifying assumptions that do not take full advantage of these data? The major difficulty with a 2-group model is choosing the number of contacts for each of the groups. The infectivities depend on the level of activity of the 2 groups. Even if the first group (the core) is 10% of the population, it is unclear how to choose a single average number of contacts for the remaining 90% of the population. Consider a 2-group core/noncore model with noncore members having 27 contacts per year (2-group model A in Table 3, Fig. 10), the actual numeric average. Note that this average includes group 6, whose members have no contacts. Surely, they should not be included, but group 5 has almost no contacts (3 per year) and almost no effect on the epidemic. Group 5 probably should not be included. Group 4 has 15 contacts per year, enough to be significantly important in the epidemic, but it plays much less of a role than Groups 2 and 3, with 81 and 33 contacts per year, respectively. Our 2-group model B instead uses a weighted average to obtain a noncore

average number of contacts of 53 per year, weighting the groups in proportion to the number of contacts they have. In fact, neither model A nor model B is perfect, which emphasizes the need for 6 groups, reflecting the SFCCC survey data.

Changing our model to a 2-group study, model A yields (0.024, 0.005, and 0.091) optimized infectivities for primary, asymptomatic, and symptomatic stages, respectively, resulting in an underestimation of symptomatic stage infectivity by approximately two thirds. (This is in keeping with results in prior 2-group studies.) Two-group study model B results in optimized infectivities of (0.028, 0.0, and 0.089). Once again, the symptomatic stage is underestimated by approximately two thirds.

Similarly, if we collapse all 6 activity groups into a single average activity level, the symptomatic stage infectivity is underestimated (0.045, 0.000, and 0.088) in the best-optimized 1-group model. The goodness-of-fit error is 0.03 compared with 0.016 for our 6-group optimization fit. Even with high symptomatic stage infectivities, a 1-group model is unable to capture the beginning of the San Francisco epidemic before

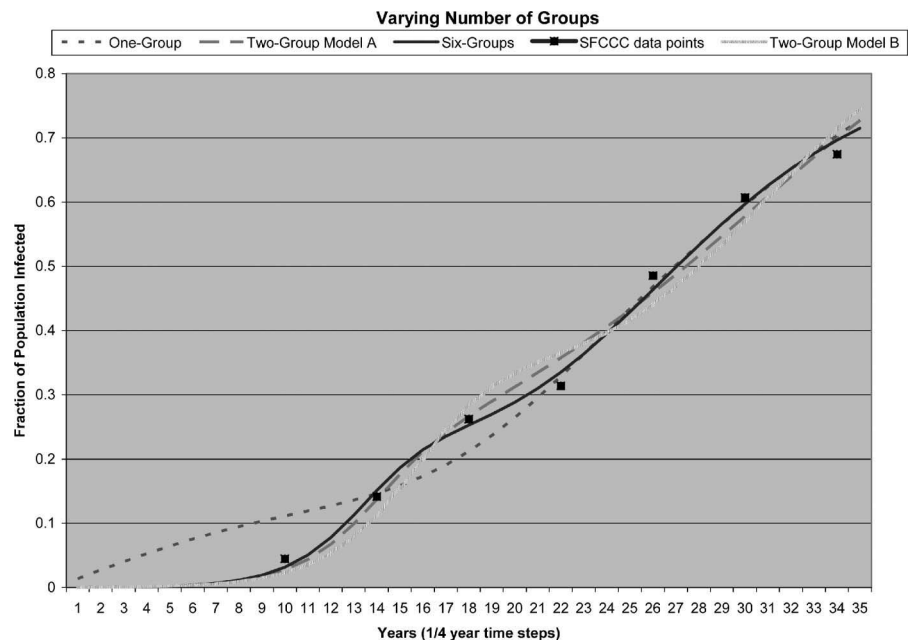


FIGURE 10. Comparison of our 6-, 2-, and 1-group models. The 1-group model cannot account for the initial growth of the epidemic. The 6-group model has the smallest square root of the average error and gives the best fit to the San Francisco City Clinics Cohort data points.

1979 (see Fig. 11). Details about the comparisons between the different group optimized models are given in Table 3.

What these comparisons show is that using fewer than 6 activity groups increasingly makes infectivity estimates artifacts of the simplifying assumptions such as core/noncore or the single activity group used in prior modeling. The 6-group models reflect the diversity of the population and reflect the reported data. What they teach us is that late in the disease, when more than half of the population is infected, lower sexual activity groups are being infected but the epidemic is accelerating. The SFCCC data show that large numbers of persons become infected, but there are fewer contacts with susceptibles because those still susceptible have few contacts. Most of these infections result from partners in the symptomatic stage. This requires a higher symptomatic stage infectivity estimate than previous models, which failed to reflect the low activity of remaining susceptibles. After 1980, the gay epidemic becomes dominated by symptomatic stage transmission.

Constant Sexual Activity Levels Assumption

The data we use is from the San Francisco epidemic up to 1984. We have assumed that each individual has a sexual activity level that does not change significantly over the period we are modeling. The Bell-Weinberg study,³⁹ a Kinsey Institute study of the 1969 San Francisco population, found that the most active 28% of men had 51 or more partners “in the past year.” The study also found that the men who were most active

over their lifetimes (again, 28%) had 1000 or more partners (Table 4). Such a total requires many years of high activity levels, perhaps 1 or 2 decades or longer. Our model is appropriate for such a population. Of course, there are some individuals who change behavior, but a large fraction of the most active “core” population remains highly active when in the symptomatic stage. The epidemic went from 4.5% to 60% infected in just 5 years.

One can hypothesize populations with variable activity levels. Some earlier models assume that individuals vary in their activity level over time. For example, Koopman et al²⁴ assume that this variation is quite rapid. They use 2 activity levels: a core with 5% of the population and the less active noncore. They assume that individuals remain in the core for an average of 1 year, stating, “Our models are not intended to reflect the transmission dynamics of any real population” (24, page 250). The assumption implies that there is virtually no correlation between the activity level of a man when he becomes infected and the activity level a few years later when he is in the symptomatic stage. (Under that assumption, the probability of an individual in the core remaining in the core for 6 years is e^{-6} or 1/400). Having people rapidly switch activity levels is quite similar to assuming that there is a single activity level.

Some may suggest, quite plausibly, that it is likely people in the symptomatic stage are less active because of the effects of HIV. If so, the symptomatic stage infectivity would have to be higher than our 0.299 to account for the large number of observed cases. This can be described by a mathematic relation. If you cut the number of contacts in half uniformly for the symptomatic stage and for all activity groups, the optimal symptomatic stage infectivity would be doubled so as not to decrease the number of new infections below the observed level.

To account for the large number of new infections in the latter stage of the epidemic, during which most of the susceptibles were from low-activity groups, one must have high

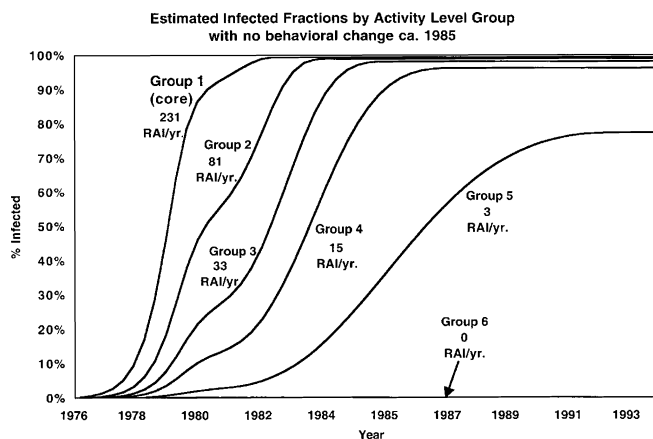


FIGURE 11. Fraction of infecteds in the bathhouse by activity group. The susceptible fraction of the entire population over time has been partitioned into susceptible fractions for each of 6 different activity groups over time. Our “bathhouse” assumption is that men vary greatly in how often they make themselves available for sex, that is, “choose partners in the bathhouse,” but once inside the bathhouse, the choice of partner is rather indiscriminate and every susceptible man’s per-contact chances of becoming infected are equal. In the bathhouse, susceptibles are similarly determined. More active persons usually get infected sooner than less active persons. The susceptible rates for more active groups drop faster than for less active groups. The graph shows, for example, that when total population susceptibles dropped to 46% (in mid-1983), only 12% of men in the bathhouse were susceptible and virtually none of the core members were susceptible (see appendix for methods).

TABLE 4. Bell-Weinburg Survey of 572 Homosexual Men in San Francisco in 1969

Sexual Partnerships (N = 572)			
No. of Homosexual Partners Ever	Percentage	No. Homosexual Partners in Past Year	Percentage
1	0	0	3
2	0	1–2	8
1–2	1	3–5	10
3–4	2	6–10	12
5–9	3	11–19	12
10–14	3	20–50	27
15–24	8	51+	28
25–49	9		
50–99	15		
100–249	17		
250–499	15		
500–999	15		
1000+	28		

From Bell AP, Weinburg S. *Homosexualities: A Study of Diversity Among Men and Women*. New York: Simon and Schuster; 1978.

symptomatic stage infectivity. Only the 6-group model is able to determine the high level of infectivity in the symptomatic stage.

Model Assumes No Decrease in Activity as the Epidemic Explodes

What if activity levels declined as the epidemic progressed? It is likely that activity levels began to decline as people became aware of some new gay disease around 1984 or 1985.⁶ We have tried alternatives to our model, for example, by cutting the contact rate in half in 1983 through 1984. We then again determined the infectivities for the 3 stages that resulted in the best fit of the data. The main effect is that the symptomatic stage infectivity must be higher than in our standard model. The infectivities of the first 2 stages are largely determined by the need to fit the pre-1981 beginning of the epidemic when there were few symptomatic stage men. Decreasing the activity while maintaining the number infected results in higher symptomatic infectivity. Our main conclusion in this report is that the symptomatic stage is far more infectious than the earlier stages. Our conclusion remains valid when there is a decrease in sexual activity level as the epidemic progresses. Our symptomatic stage estimate, although higher than what prior studies report,^{6,16,24} is, in fact, a lower limit. Indeed, Hethcote and Van Ark's assumption of decreased activity beginning in 1981 requires them to use an AIDS stage infectivity of 0.75 to model the SFCCC data.¹⁴

Stochastic Versus Deterministic Models

Our model is a deterministic model. At each time step, we determine the fraction of individuals who make the transition from a given stage of infection to the next for each activity level, such as from core susceptible to core primary infection. We also developed a stochastic model in which, at each time step, we first compute the fraction of people who would make each transition according to the rules of our deterministic model and then convert this to a number of men by considering San Francisco gay population to be an estimated 70,000 individuals from the late 1970s through early 1980s.²⁸ Using this fraction as a mean, we select a random number from a Poisson distribution. This random number becomes the number of individuals who make the transition at that time. The epidemic is then simulated, repeating this Poisson process for each time step and for every transition (see appendix). We find that the biggest differences between the epidemics of the deterministic model and the stochastic model occur when the fraction of people infected is small. By the time the fraction of infected individuals reaches 4.5% (as was the case in 1978, the SFCCC first data point), both models generate similar curves.

Because we are using the same 3 infectivities in both models, our confidence in these 3 infectivities increases. When their results are congruent, there is negligible harm in using a deterministic model, even when a stochastic model might seem more appropriate theoretically. Both approaches require high levels of symptomatic stage infectivity compared with primary stage infectivity.

We conclude that as compared with the assumptions used by other researchers, ours are more realistic and/or appropriate when applied to the SFCCC population. Our finding that symp-

tomatic stage infectivities are approximately 30 times higher than previous estimates is not an artifact of our assumptions.

DISCUSSION OF MODELING CONCLUSIONS

Testing Whether the Conclusion Symptomatic Infectivity Is Greater Than Primary Infectivity

To test this conclusion further, we examine the cases in which the 2 infectivities are equal. For comparison, we ran our optimization code with the added constraint that the primary stage infectivity equals the symptomatic stage infectivity. The best-fit infectivities are 0.0, 0.02, and 0.0, respectively, for the 3 stages, and the RMS error is 0.036. Note that the RMS error is more than twice the RMS error for the best-fit infectivities mentioned previously. With this constraint, there is no best fit with positive values for the primary and symptomatic stage infectivities.

Interval of Infectivity Estimates

Our infectivity results must be valid for variations in model parameters. § Previously, we discussed error in the data points and the effect that variations in the data would have on our infectivity results. We also mentioned the differences between stochastic and deterministic models. Infectivities would increase if we considered a behavior change before 1985 because of knowledge of the disease or associated with progression of the disease. We could also take into account the possibility of preferential mixing, in which one chooses partners within their own group. If one does not allow those in lower activity groups to have contacts with those in the core and other more active groups, the infectivity of the third stage would have to be even higher than our estimates for the epidemic to rip through the 6 groups as shown in Figure 11. This conclusion holds when one considers a combination of the bathhouse mixing and preferential mixing.

Allowing all combinations of the parameters mentioned previously, we end up with intervals of infectivity estimates. The primary stage varies from 0.014 to 0.024, the asymptomatic stage varies from 0.000 to 0.008,** and the symptomatic stage varies from 0.126 to 0.493. We are interested in the ratio of the symptomatic stage infectivity to the primary stage infectivity, which varies from 8.6 to 33.7. Thus, we conclude that even under variations of our model parameters, the symptomatic stage remains significantly more infectious than the primary stage.

CONCLUSION

Earlier models shed little light on slow transmission epidemics such as the African and San Francisco epidemics from 1980. They are dominated by symptomatic stage transmission, and they seriously distort the transmission

§Our model has a time step of 0.25 year, 2 substages for the primary period, 3 substages for the asymptomatic period, and 1 substage for the symptomatic period. We also considered time steps of 0.33 and 0.5 year, along with 1 to 2 substages for the primary stage, 1 to 6 substages for the asymptomatic stage, and 1 to 2 substages for the symptomatic stage.

**If an infectivity value becomes negative in our optimization routine, we set it to 0. We do not believe that the asymptomatic period is ever truly 0.

dynamics after 1980. In 2-wave epidemics such as the San Francisco gay epidemic, there is a period when primary infection stage transmission is the predominant mode of transmission and alone can sustain the epidemic. Only symptomatic stage infection can sustain a slow epidemic such as the epidemic in Africa. If there is no such period, you get an epidemic such as that in South Africa. Ultimately, both epidemic patterns become dominated by symptomatic stage transmission.

Underestimating the symptomatic stage infectivity results in a severe underestimation of R_0 , the severity of the epidemic, and the measures necessary to end the epidemic. Our results provide a firm basis for a needed systematic reassessment of prevailing wisdom and strategies concerning the control, containment, and management of the HIV pandemic. Our results imply that screening of at-risk populations can identify most infected individuals before they enter their most infectious stage. Removal of symptomatic stage transmission would reduce R_0 to less than 1 for many extant at-risk populations (although not the SFCCC population).

REFERENCES

- Moss AR, Osmond D, Bacchetti P, et al. Risk factors for AIDS and HIV seropositivity in homosexual men. *Am J Epidemiol*. 1987;125:1035–1047.
- Ostrow DG, DiFranceisco WJ, Chmiel JS, et al. A case-control study of human immunodeficiency virus Type 1 seroconversion and risk-related behaviors in the Chicago MACS/CCS Cohort, 1984–1992. *Am J Epidemiol*. 1995;142:875–883.
- Centers for Disease Control and Prevention. Update: acquired immunodeficiency syndrome in the San Francisco cohort study, 1978–1985. *MMWR*. 1985;34:573–575.
- Jaffe HW, Darrow WW, Echenberg DF, et al. The acquired immunodeficiency syndrome in a cohort of homosexual men: a six-year follow-up study. *Ann Intern Med*. 1985;103:210–214.
- Curran JW, Morgan WM, Hardy AM, et al. The epidemiology of AIDS: current status and future prospects. *Science*. 1985;229:1352–1357.
- Ahlgren DJ, Gorny MK, Stein AC. Model-based optimization of infectivity parameters: a study of the early epidemic in San Francisco. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1990;3:631–643.
- Pantaleo G, Graziosi C, Fauci AS. Review article: the immunopathogenesis of human immunodeficiency virus infection. *N Engl J Med*. 1993;328:327–335.
- Clark SJ, Saag MS, Decker WD, et al. High titers of cytopathic virus in plasma of patients with symptomatic primary HIV infection. *N Engl J Med*. 1991;324:954–960.
- Darr ES, Moudgil T, Meyer RD, et al. Transient high-levels of viremia in patients with primary human immunodeficiency virus type 1 infection. *N Engl J Med*. 1991;324:961–964.
- Anderson RM. The spread of HIV and sexual mixing patterns. In: Mann J, Tarantola D, eds. *AIDS in the World II: Global Dimensions, Social Roots, and Responses*. The Global AIDS Policy Coalition. New York: Oxford University Press; 1996:71–86.
- Zhang H, Dornadula G, Beumont M, et al. Human immunodeficiency virus type 1 in the semen of men receiving highly active antiretroviral therapy. *N Engl J Med*. 1998;339:1803–1809.
- Bernard EJ. HAART and HIV transmission. *AIDS Treatment Update* 2002;118:2–5.
- Gupta P, Mellors J, Kingsley L, et al. High viral load in semen of human immunodeficiency virus type 1-infected men at all stages of disease and its reduction by therapy with protease and nonnucleoside reverse transcriptase inhibitors. *J Virol*. 1997;71:6271–6275.
- Hethcote HW, Van Ark JW. *Modeling HIV Transmission and AIDS in the United States*. Berlin: New York, Springer-Verlag; 1992.
- Longini IM, Jr, Clark WS, Haber M, et al. The stages of HIV infection: waiting times and infection transmission probabilities. In: Castillo-Chavez C, ed. *Mathematical and Statistical Approaches to AIDS Epidemiology*. Berlin: New York, Springer-Verlag; 1989:110–137.
- Jacquez JA, Koopman JS, Simon CP, et al. Role of the primary infection in epidemics of HIV infection in gay cohorts. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1994;7:1169–1184.
- Grant RM, Wiley JA. Infectivity of the human immunodeficiency virus: estimates from a prospective study of homosexual men. *J Infect Dis*. 1987;156:189–193.
- Padian N, Shiboski SC, Glass SO, et al. Heterosexual transmission of human immunodeficiency virus (HIV) in northern California: results from a ten-year study. *Am J Epidemiol*. 1997;146:350–357.
- DeGruttola V, Searge GS III, Mayer KH, et al. Infectiousness of HIV between male homosexual partners. *J Clin Epidemiol*. 1989;42:849–856.
- Peterman TA, Stoneburner RL, Allen JR, et al. Risk of human immunodeficiency virus transmission from heterosexual adults with transfusion-associated infections. *JAMA*. 1988;259:55–58.
- Hethcote HW, Yorke JA. *Gonorrhea Transmission Dynamics and Control*. Berlin: New York, Springer-Verlag; 1984.
- Doll LS, Darrow W, O'Malley P, et al. Self-reported behavioral change in homosexual men in the San Francisco City Clinics Cohort. Presented at the III International Conference on AIDS, Washington, June 1987.
- Centers for Disease Control and Prevention. Self-reported changes in sexual behaviors among gay and bisexual men—San Francisco. *MMWR*. 1985;34:613–615.
- Koopman JS, Jacquez JA, Welch GW, et al. The role of early HIV infection in the spread of HIV through populations. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1997;14:249–258.
- Winkelstein W, Lyman DM, Padian N, et al. Sexual practices and risk of infection by the human immunodeficiency virus: the San Francisco Men's Health Study. *JAMA*. 1987;257:321–325.
- Werdegard D, O'Malley P, Bodecker T, et al. Self-reported changes in sexual behaviors among homosexual and bisexual men from the San Francisco City Clinics Cohort. *MMWR*. 1987;36:187–189.
- Kaplan EH. Modeling HIV infectivity: must sex acts be counted? *J Acquir Immune Defic Syndr*. 1990;3:55–61.
- UNAIDS—Joint United Nations Programme on HIV/AIDS. The evolving picture by region. In: Report on the Global HIV/AIDS Epidemic. UNAIDS/World Health Organization; 1998.
- Gray R, Wawer MJ, Brookmeyer R, et al, and the Rakai Project Team. Probability of HIV-1 transmission per coital act in monogamous heterosexual, HIV-1-discordant couples in Rakai, Uganda. *Lancet*. 2001;357:1149–1153.
- Saracco A, Musicco M, Nicolosi A, et al. Man-to-woman sexual transmission of HIV: longitudinal study of 343 steady partners of infected men. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1993;6:497–502.
- Piatok M, Jr, Saag MS, Yang LC, et al. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science*. 1993;259:1749–1754.
- De Vincenzi I. A longitudinal study of human immunodeficiency virus transmission by heterosexual partners. *N Engl J Med*. 1994;331:341–346.
- Kanabus A, Allen S. The origins of AIDS and HIV and the first cases of AIDS. Available at: <http://www.avert.org/origins.htm>.
- Nahmias AJ, Weiss J, Yao X, et al. Evidence for human infection with an HTLV/LAV-like virus in Central Africa. *Lancet* 1986;i:1279–1280.
- Mann J, Tarantola D. Global overview: a powerful HIV/AIDS pandemic. In: Mann J, Tarantola D, eds. *AIDS in the World II: Global Dimensions, Social Roots, and Responses*. The Global AIDS Policy Coalition. New York: Oxford University Press; 1996:5–40.
- Centers for Disease Control and Prevention. HIV incidence among young men who have sex with men—seven U.S. cities, 1994–2000. *MMWR*. 2001;50:440–444.
- Pinkerton SD, Abramson PR. Implications of increased infectivity in early-stage HIV infection: application of a Bernoulli-process model of HIV infection. *Eval Rev*. 1996;20:516–540.
- Blythe SP, Anderson RM. Variable infectiousness in HIV transmission models. *IMA J Math Appl Med Biol*. 1988;5:181–200.
- Bell AP, Weinburg S. *Homosexualities: A Study of Diversity Among Men and Women*. New York: Simon and Schuster; 1978.
- Blower SM, Gershengorn HB, Grant RM. A tale of two futures: HIV and antiretroviral therapy in San Francisco. *Science*. 2000;287:650–654.
- Saltzman SP, Stoddard AM, McClusker J, et al. Reliability of self-reported sexual behavior risk factors for HIV infection in homosexual men. *Public Health Rep*. 1987;102:692–697.
- Pinkerton SD, Abramson PR. Occasional condom use and HIV risk reduction. *J Acquir Immune Defic Hum Retrovirol*. 1996;13:456–460.

43. Palacino H. Safer sex. In: Cohen PT, Sande MA, Volberding A, eds. The AIDS Knowledge Base.
44. Yorke JA, Nathanson N, Pianigiani G, et al. Seasonality and the requirements for perpetuation and eradication of viruses in populations. *Am J Epidemiol.* 1979;109:103–123.
45. Centers for Disease Control and Prevention. Current trends update: impact of the expanded AIDS surveillance case definition for adolescents and adults on case reporting—United States. *MMWR.* 1993;43:160–161, 167–170.

APPENDIX

Stage Model

For primary infection, the substages amount to the assumption that the average time after infection before one infects others is a 0.25 year. The peak of the viral load occurs at 0.25 year.¹⁰

We assume that susceptibles are in 3-month cohorts moving through susceptible and HIV infection stages as determined by the following variables and equations, where *i*, *j* = activity groups 1 (core) and 2 through 6; *k* = primary, asymptomatic, symptomatic; and *t* = 1, 2, 3, We assume that primary infection lasts 2 time periods but that all primary stage transmission occurs at the end of the first time period (ie, 0.25 year). We assume that new susceptibles enter the population at the same rate at which there are AIDS deaths.

We now define the variables used in our model:

- Susceptible fraction of group *i* at time *t*: S_i^t
- Primary first fraction of group *i* at time *t*: $P_{1,i}^t$
- Primary second fraction of group *i* at time *t*: $P_{2,i}^t$
- Asymptomatic first fraction of group *i* at time *t*: $L_{1,i}^t$
- Asymptomatic second fraction of group *i* at time *t*: $L_{2,i}^t$
- Asymptomatic third fraction of group *i* at time *t*: $L_{3,i}^t$
- Symptomatic fraction of group *i* at time *t*: A_i^t
- Death fraction of group *i* at time *t*: D_i^t
- Rate of partnering of group *i* with group *j*: r_{ij}
- Infectivity of persons at stage *k*: a_k
- Fraction of group *j* that is in stage *k* at time *t*: f_{jk}^t
- Duration of asymptomatic substages: *d*
- Time step (in terms of fraction of a year): Δt

$$P_{1,i}^{t+1} = S_i^t \left[\sum_j \sum_k r_{ij} a_k f_{jk}^t \right] \text{ (new group i primary fraction 1)}$$

$$P_{2,i}^{t+1} = P_{1,i}^t \text{ (new group i primary fraction 2)}$$

$$L_{1,i}^{t+1} = L_{1,i}^t + P_{2,i}^t - \frac{L_{1,i}^t \Delta t}{d} \text{ (new group i asymptomatic fraction 1)}$$

$$L_{2,i}^{t+1} = L_{2,i}^t + \frac{L_{1,i}^t \Delta t}{d} - \frac{L_{2,i}^t \Delta t}{d} \text{ (new group i asymptomatic fraction 2)}$$

$$L_{3,i}^{t+1} = L_{3,i}^t + \frac{L_{2,i}^t \Delta t}{d} - \frac{L_{3,i}^t \Delta t}{d} \text{ (new group i asymptomatic fraction 3)}$$

$$A_i^{t+1} = A_i^t + \frac{L_{3,i}^t \Delta t}{d} - \frac{A_i^t \Delta t}{3} \text{ (new group i symptomatic fraction)}$$

$$D_i^{t+1} = \frac{A_i^t \Delta t}{3} \text{ (new group i death fraction)}$$

$$S_i^{t+1} = S_i^t - P_{1,i}^{t+1} + D_i^{t+1} \text{ (new group i susceptible fraction)}$$

From this model we can derive:

$$I^t = \sum_{i=1}^6 \left(2 * P_{1,i}^t + L_{1,i}^t + L_{2,i}^t + L_{3,i}^t + A_i^t \right) \text{ (population cumulative infected fraction)}$$

$$C^t = \sum_{i=1}^6 P_{1,i}^t \text{ (population seroconversion rate)}$$

Stage Transitions

It is well established that in the first approximation, a typical untreated HIV infection progresses through the stages defined by successively high, low, and then very high viral levels in the blood.⁶ For the SFCCC study, the average durations of each stage are well documented and confirmed by modeling.⁶ The problem is that those averages are based on some people who pass through stages to AIDS faster and others who pass through slower or perhaps never become symptomatic or develop AIDS. This is only an issue for the relatively long asymptomatic and symptomatic AIDS stages. A good model needs to have some sort of diffusion pattern producing different rates of passage through those stages. Because there are no data regarding such “diffusion,” the simplest way to model this is to allow substages, where the number of substages determines the diffusion gradient. One can then choose how many substages to use depending on accuracy of best fit to HIV incidence with different numbers of substages. Our basic model does this and has 3 diffusion substages for the asymptomatic period and 1 stage for symptomatic/AIDS (see Fig. 5).

Model Parameters

Most model parameter values are specified using SFCCC data. Average time from seroconversion (development of antibodies) to death was reportedly 10.3 years.^{††} In our model, the average primary infectious period lasts for 0.5 year, although seroconversion typically occurs at approximately 3 months. When we tried shorter primary stage periods such as 0.33 year, we were unable to fit the data as well. The epidemic behaves as if semen remains infectious for a bit longer than the usual primary stage period. Thus, the average duration of an HIV infection is 10.5 years. The remaining parameters are determined by best-fit approximation to the 1978 through 1984 SFCCC HIV infection growth data. Thus, every parameter value in the stage model is a firm SFCCC datum or is highly constrained by the SFCCC data. No unconstrained parameter assumptions are used.

^{††}Individuals undergoing antiretroviral treatment tend to progress to AIDS at a much slower rate than untreated individuals.⁴

Interpreting the Stage Model

The San Francisco population is divided into 6 activity groups, and each group is divided into 3 stages of infection (primary, asymptomatic, and symptomatic). To run our model, the user specifies the fraction of the men in each stage at an initial time, t_0 , and the 3 infectivities for the 3 stages of infection.

Given the fraction of each activity group that is in each stage at time t , the rules built into the model dictate what the corresponding fractions will be at time $t + \Delta t$, where Δt is a specified fraction of a year. We typically took Δt to be 0.25 or 0.33 year for the time step, and we report results here for 0.25 year, although the results for 0.33 year are similar. The model takes these fractions and takes another time step, applying the same procedure to obtain the corresponding fractions for time $t + 2 \Delta t$. The model takes a certain specified number of steps long enough for it to create a record of an outbreak similar to San Francisco's.

Model's Bookkeeping of New Infections in 1 Time Step

Given the fraction in each stage of each activity group, the model computes the expected number of contacts for all the men in each of the 4 stages: N_{sus} , N_p , N_a , and N_s . The risk, R , of a susceptible man becoming infected from 1 contact is:

$$R = \frac{(I_{sus} * N_{sus} + I_p * N_p + I_a * N_a + I_s * I_a)}{(N_{sus} + N_p + N_a + N_s)}$$

In each activity group, the fraction of men newly infected, F_{new} , at time $t + \Delta t$ is the fraction susceptible times the number of contacts each man has in time Δt times the risk, R . To obtain the susceptible population for each activity group for the time $t + \Delta t$, we subtract the fraction F_{new} from the susceptible fraction for time t and add F_{new} to the primary stage for time $t + \Delta t$.

Model's Bookkeeping of the Fraction in Each Stage

If the average duration for a stage is Y years, the fraction $\Delta t/Y$ of people in that stage is moved to the next stage. We use $2\Delta t$ years for the duration of primary stage, 7 years for the duration of asymptomatic stage, and 3 years for the duration of symptomatic stage.

It should be noted that the real meaning of the $2\Delta t$ year primary period (which seems rather long to us) is that people who are initially infected at time t can create new infections at time $t + \Delta t$ and the number of contacts they have while in the primary stage is $2\Delta t$ times the number of contacts for a year. One could alternatively say that the primary period is Δt and double the infectivity for the period.

Initializing the Model

We do not know when the epidemic actually began in San Francisco nor do we know the initial state, which really does not matter. No matter how we initialize the outbreak (whether the initial man or men were highly active or less active or in the primary stage or the symptomatic stage), we must choose an initial time so that the epidemic reaches the prevalence (ie, the fraction infected) of 4.5% in 1978. By that time and thereafter, the distribution of infected people is essentially independent of how we started the epidemic. As long as the initial infected fraction is small, the long-term shape of the

plot of prevalence is not affected. The first prevalence report from the SFCCC study was 4.5% in 1978.

By our estimate, the 4.5% prevalence figure might better be reported as $4.5\% \pm 1.3\%$, representing an error of 1 standard deviation. The method for computing prevalence for other years is reported in less detail, and we cannot compute standard deviations for those years.

Optimization Procedures

Optimization considerations determined the choice of 3 equal-duration asymptomatic substages. There are 6 activity level groups (see Table 2). Fractions of susceptibles who become infected in a given period are estimated using products of stage infectivities, fractions of groups in a given stage, and partnering rates between groups. Passage from a stage or substage to the next is proportional to average stage or substage duration. Using SFCCC data for other model parameters, stage infectivities are estimated by optimized fits to SFCCC HIV incidence data (see Table 1). An optimization routine (Newton's method applied to the gradients of $E = \text{root-mean-square errors in SFCCC fit to data for the years 1978 to 1984}$) was used to obtain best-fit approximation to SFCCC data estimates for average primary, asymptomatic, and symptomatic stage infectivities. This produced the solution in Figure 6.

Sensitivity Analysis

Although the SFCCC HIV incidence data are of unusually high resolution and are biologically based, they are still subject to measurement and sampling errors. Furthermore, the behavioral data are self-reported. Studies indicate that such self-reported behaviors among gay men are quite reliable.⁴¹⁻⁴³ Nevertheless, there are errors inherent in such data.

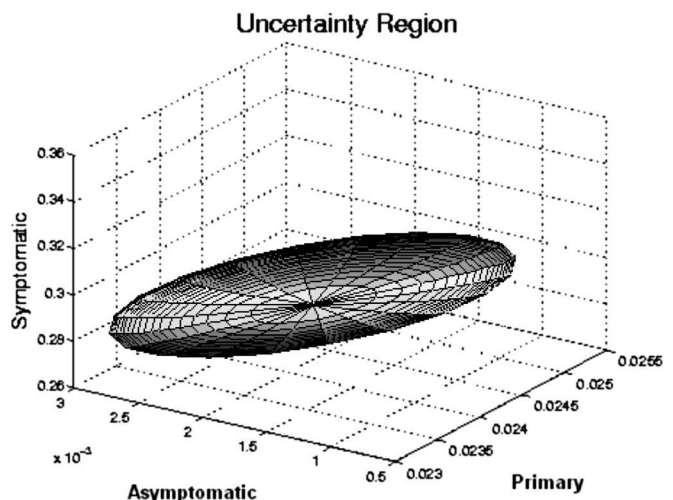


FIGURE 12. Sensitivity analysis. The 3 infectivities for our 3 stages depend on the reported epidemic data for years 1978 through 1984. There is an inherent error in the data, perhaps 1% to 2%. If these data points are changed by a (root-mean-square) average of 2%, the infectivities lie on the ellipsoid shown. If they are changed by 4%, the size of the ellipse is doubled in each direction. If they are not changed, they lie at the center of the ellipsoid (see appendix for methods).

When we vary the fraction of the population infected for the data points 1978 through 1984 by approximately ± 0.02 and then apply these optimization procedures, we get slightly different infectivities for our 3 stages. In summary, primary stage infectivity varies in the range of 2.3% to 2.5%, asymptomatic stage infectivity varies in the range of 0.0% to 0.4%, and symptomatic/AIDS stage infectivity varies in the range of 25.4% to 34.4%. Figure 12 shows us the region of uncertainty.

Fraction of Infecteds in the Bathhouse by Activity Group

From the model equations, we calculate the proportion of infecteds in each activity group versus time as follows:

$$I_i^t = 2 * P_{1,i}^t + L_{1,i}^t + L_{2,i}^t + L_{3,i}^t + A_i^t \quad \text{for } i = 1, 2, 3, 4, 5, 6.$$

The results are shown in Figure 11.

Stochastic Model

As mentioned previously, we developed a stochastic model in which, at each time step, we first compute the fraction of people who would make each transition according to the rules of our deterministic model and then convert this to a number of men by considering the San Francisco gay population in the late 1970s and early 1980s to be an estimated 70,000 individuals.²⁸ Using this fraction as a mean, we select a random number from a Poisson distribution. This random number becomes the number of individuals who make the transition at that time. The epidemic is then simulated repeating this Poisson process for each time step and for every transition.

Let f_i be the fraction of population in activity group I , and T be the total population ($N = 70,000$).

1. Stochastic model
 - 1.1. $P_{1,i}^{t+1} = P_{1,i}^t$ (fraction that transitions from P1 to P2 at time t)
 - 1.2. Let $\lambda = P_{1,i}^t * T$
 - 1.3. Compute random variable from 0 to 1, r .
 - 1.4. Find N such that:

$$\sum_{n=0}^N \frac{e^{-\lambda} \lambda^n}{n!} = r$$

- 1.5. Let:

$$P_{1,i}^{t+1} = N / f_i / T$$

2. Repeat procedure

$$P_{2,i}^{t+1} = P_{2,i}^t \quad \text{(fraction that transition from P2 to L1 at time } t)$$

$$L_{1,i}^{t+1} = \frac{L_{1,i}^t}{7} \quad \text{(fraction that transition from L1 to L2 at time } t)$$

$$L_{2,i}^{t+1} = \frac{L_{2,i}^t}{7} \quad \text{(fraction that transition from L2 to L3 at time } t)$$

$$L_{3,i}^{t+1} = \frac{L_{3,i}^t}{7} \quad \text{(fraction that transition from L3 to A at time } t)$$

$$A^{t+1} = \frac{A^t}{3} \quad \text{(fraction that transition from A to death at time } t)$$