Numerical Results of the Risk-Based Model

by James M. Hyman, E. Ann Stanley, and Stirling A. Colgate

Here we will present numerical solutions to the full risk-based biased-mixing model. These solutions validate the simplified version of the model presented in the main text and illustrate how variations in the input parameters affect the predicted course of the epidemic. The equations and parameters of the model are defined in "Mathematical Formalism for the Risk-Based Model of AIDS," hereafter referred to as "Math Formalism." The model tracks the time evolution of three sectors of the population: the sexually active susceptible $S(t, r)$; the sexually active infecteds $I(t, \tau, r)$; and the people with AIDS $A(t, \tau, r)$. It takes into account deaths due to AIDS and the long time between HIV infection and conversion to AIDS. It also allows us to vary assumptions about the infectiousness as a function of time since infection and the mixing between various risk groups in the population.

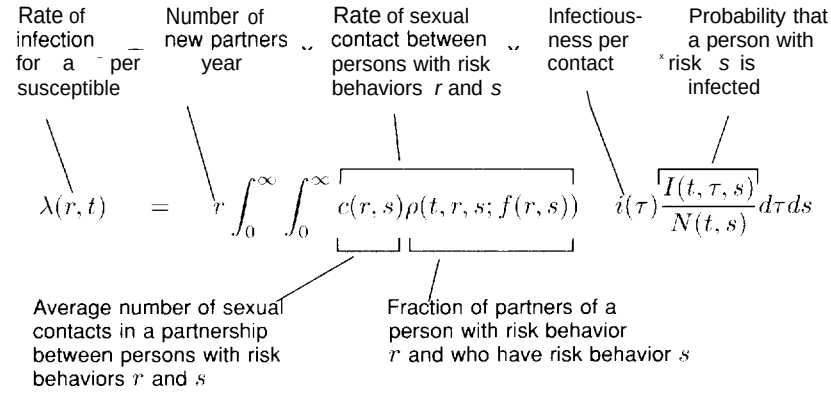
First we will assess the validity of the predictions in the main text. The analytic calculation presented them predicted that biased mixing among the sexually active population gives rise to a saturation wave of infection, which yields power-law growth in both the number infected and the number of people with AIDS. That calculation was based on the following assumptions: the initial susceptible population $S_o(r)$ is distributed in risk behavior as r^{-3} for r greater than the mean value of r; the infectiousness i is constant; the cumulative probability of conversion to AIDS $C(\tau)$ is zero for the first two years after infection and then increases linearly with τ at a rate such that every infected individual develops AIDS by 18 years after infection; and finally, the same fraction is infected in all risk groups before the start of the saturation wave, The wave of infection was then calculated as if each risk group had a growth rate proportional to r and grew to saturation independently of all other groups. That is, we did not account for mixing between people with different risk behavior because the calculation is too difficult to perform analytically, Moreover, AIDS cases and deaths were not removed from the infected population. The result was that the number infected grows as t^2 and the number of people with AIDS grows at t^3.

To check whether mixing among individuals with different risk behavior alters that result, we solved the full set of equations given in "Math Formalism." We used the same assumptions and conditions outlined above except that we allowed mixing between people with different risk behavior r. We found that when mixing is restricted to people whose risk behaviors are within a factor of 2 of each other, that is, the mixing is biased, a saturation wave of infection moves from high- to low-risk groups and the number infected grows as t^2, as predicted by the analytic calculation in the main text. Also, when mixing is random, or homogeneous, that is, is based only on availability, the number infected grows exponentially, the relative growth rate is constant, and the fastest growth occurs in the population with the most likely risk. Thus, doubling times for biased mixing are shorter initially and later become longer than those for random mixing.

Now let's consider numerical solutions to the full model under more general assumptions, We will first comment on their overall behavior and then present specific solutions. The numer-

THE RATE OF INFECTION $\lambda(r,t)$

The heart of the risk-based model is the complicated functional form of the rate of infection per susceptible with risk r, $\lambda(r,t)$ (see Eqs. 10 and 13 in "Math Formalism"). We will describe this function in words:

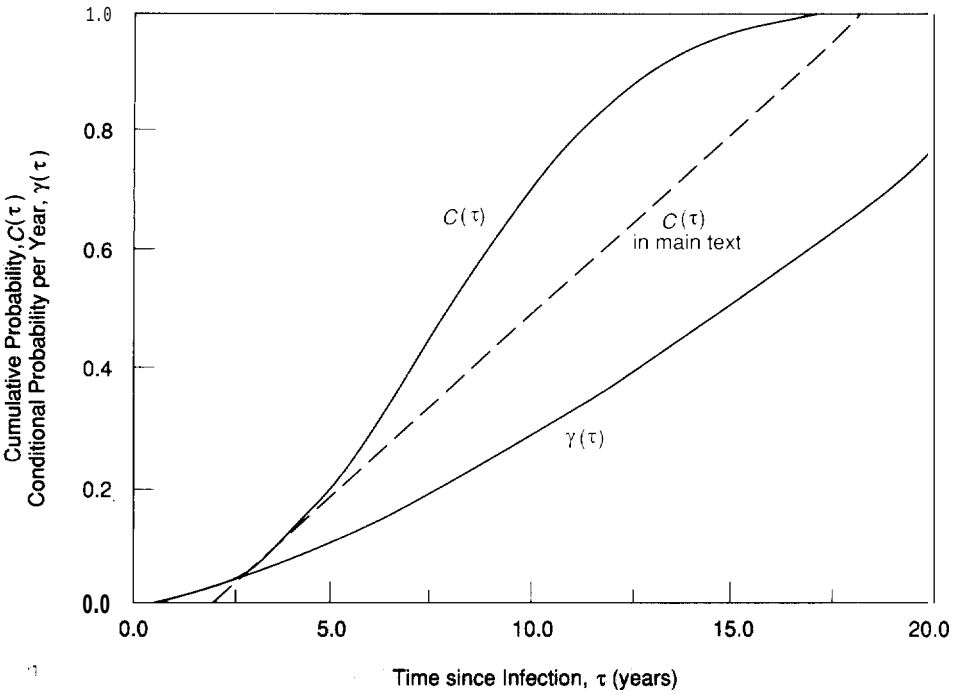


The function $\rho(t, r, s)$ describes the level of mixing between people with risk behaviors r and s. It is defined in terms of an acceptance function $f(r, s)$ that determines the range from which partners are chosen.

ical results of the model change as we vary the input parameters $S_o(r)$, $I(t,0,r)$, $i(\tau)$, $A(t,0,r)$, $c(r,s)$, $f(r,s)$, $\gamma(\tau)$, μ, and $\delta(\tau)$ (see Fig. 1 in "Math Formalism" for the definitions of these parameters). The most critical parameters for determining the course of the epidemic are the initial distribution of risk behavior among the susceptible population $So(r)$ and the functions $i(\tau)$, $c(r,s)$, and $f(r,s)$, which determine the rate of infection per susceptible $\lambda(r,t)$ (see "The Rate of Infection"). In particular, the acceptance function $f(r,$ s) specifies the amount of mixing between different risk groups. Provided the mixing is biased, $So(r)$ decays as r^{-3} or r^{-4} and the numerical value of the product $c(r,s)i(\tau)$ is between 0.025 and 0.001 (this last provision determines the time scale of the epidemic), numerid solutions of our model show that the infection travels as a saturation wave from high- to low-risk groups for approximately the first 20 years. During those years the cumulative number infected and the cumulative number of people with AIDS grow as polynomials in time, rather than as exponential.

By varying the functional forms of $\gamma(\tau)$, the rate, or conditional probability, of developing AIDS, and $i(\tau)$, the infectiousness since time of infection, we can raise or lower the degree of the polynomial growth, but in all of our calculations with biased mixing, the growth remains polynomial after the initial transients.

With these general remarks as background, we present various numerical solutions to the model. To obtain these solutions, Eqs. 9-10 in "Math Formalism" were integrated numerically with an explicit Adams-Bashford-Moulton solution method to an accuracy of 10^{-6} per unit time. The dependences on τ and r were calculated on a uniform grid of between 71 and 201 mesh points, and the convergence of solutions has been verified to within a few per cent.



RATE OF CONVERSION TO AIDS

Fig. 1. The rate of conversion to AIDS at time τ after infection $\gamma(\tau)$ is equal to the conditional probability that a person who did not have AIDS before time τ develops AIDS at time τ. Thus, it is given by $\gamma(\tau) = \frac{dC(\tau)/d\tau}{1-C(\tau)}$, where $C(\tau)$ is the cumulative probability of developing AIDS at τ years after infection. The figure shows plots of the functions $\gamma(\tau)$ and $C(\tau)$ used in all the numerical solutions presented here. For comparison we also show a plot of the form for $C(\tau)$ assumed in the main text (dashed line).

We emphasize, however, that although the solution techniques are accurate, the equations are still crude approximations and the results are meant to illustrate the general behavior of the model, not to give accurate forecasts of the future. Even the full model is much too simplistic to be used as a predictive tool.

For all the solutions presented here, we assume an initial population of 10 million people whose risk behavior (which we identify as the number of new partners per year) is distributed as an inverse cubic with a mean of 24 partners per year. We use the initial distribution $S_0(r) = 20(1 + \frac{r}{24})^{-3}$. We also use that form of $\gamma(\tau)$, the conditional probability for "converting to AIDS, shown in, Fig. 1. (The relationship between $\gamma(\tau)$ and $C(\tau)$ is described in the figure caption.) We use the constant value $\mu = 0.02$ per year for the fractional rate of maturation. The fractional rate of deaths due to AIDS $\delta(\tau)$ is obtained from CDC data. Also, for simplicity in this series of calculations, we assume the number of contacts per partner $c(r,s)$ is a constant $\bar{c}$.

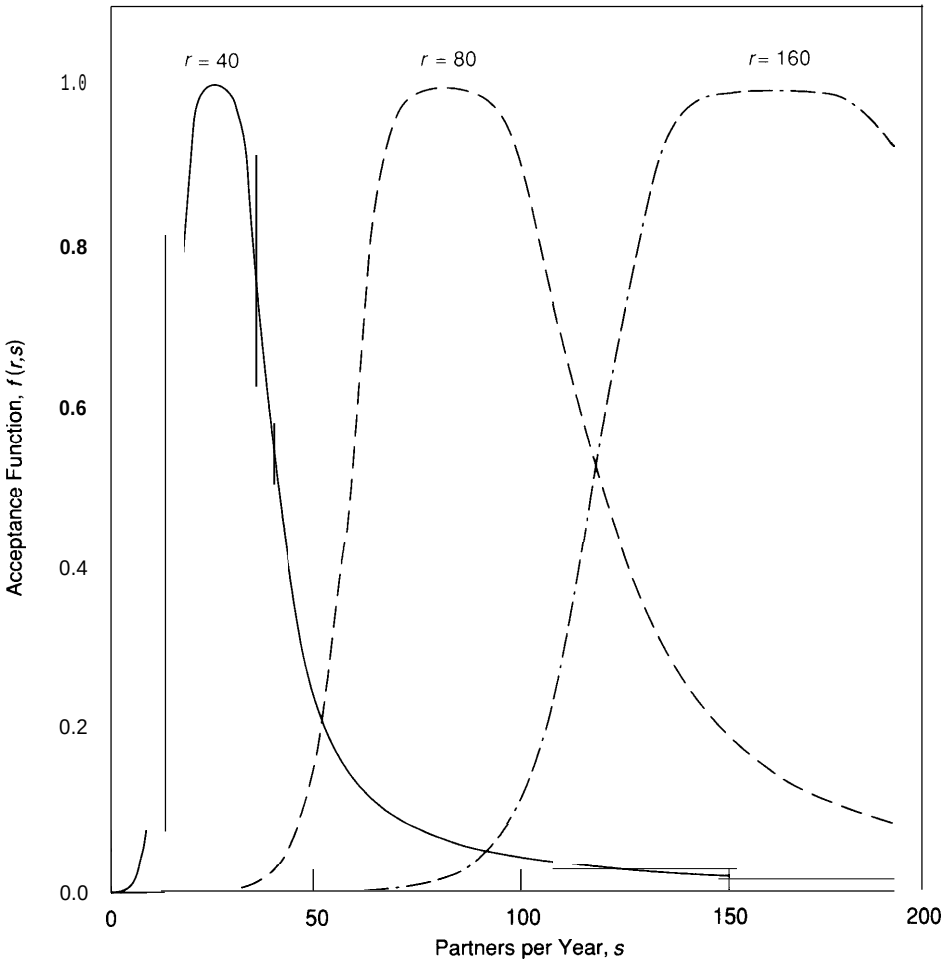
The parameter that we vary from one solution to another is $\lambda(r,t)$, the relative growth rate of infection among susceptible with r partners per year. In particular we vary two factors in $\lambda(r,t)$: the acceptance function $f(r,$ s) and the in-

BIASED MIXING FOR BASELINE SOLUTION

Fig. 2. The numerical solutions presented here use an inverse quartic function for the acceptance function $f(r, s)$:

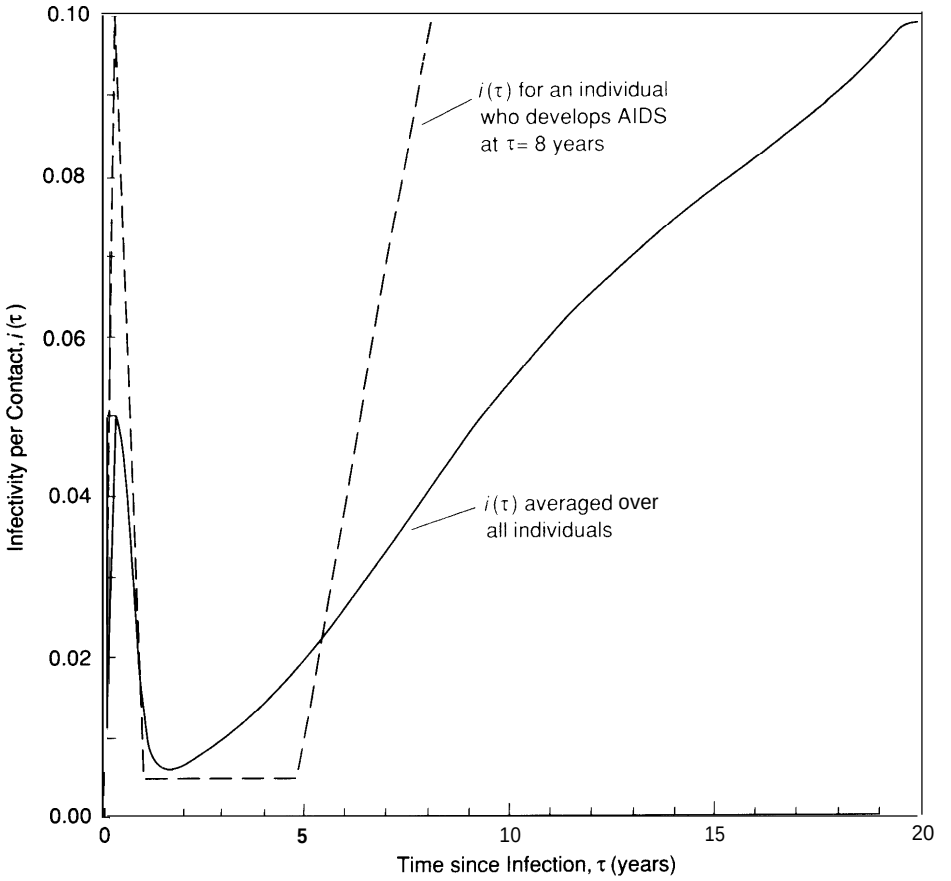
$$f(r, s) = \left[1 + \frac{(r - s)^4}{\epsilon(r + r_m)^4}\right]^{-1}.$$

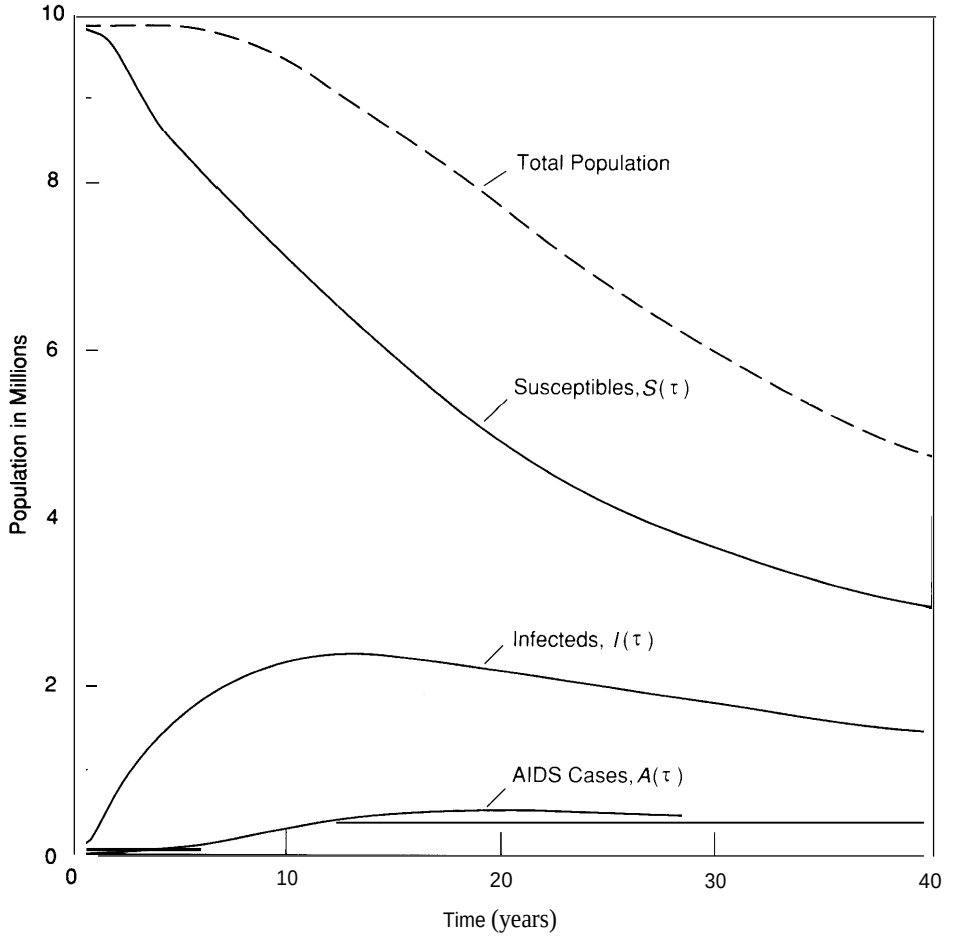
The figure shows $f(r, s)$ versus s for r = 40, 80, and 150 when ϵ = 0.01. For each value of r, $f(r, s)$ determines the fraction of partners with risks chosen by people with risk r. Here $f(r, s)$ specifies that most partners of a person with risk r have risk behaviors between ½r and r; that is, the mixing is heavily biased toward people with similar risk behavior.



TIME-DEPENDENT INFECTIVITY

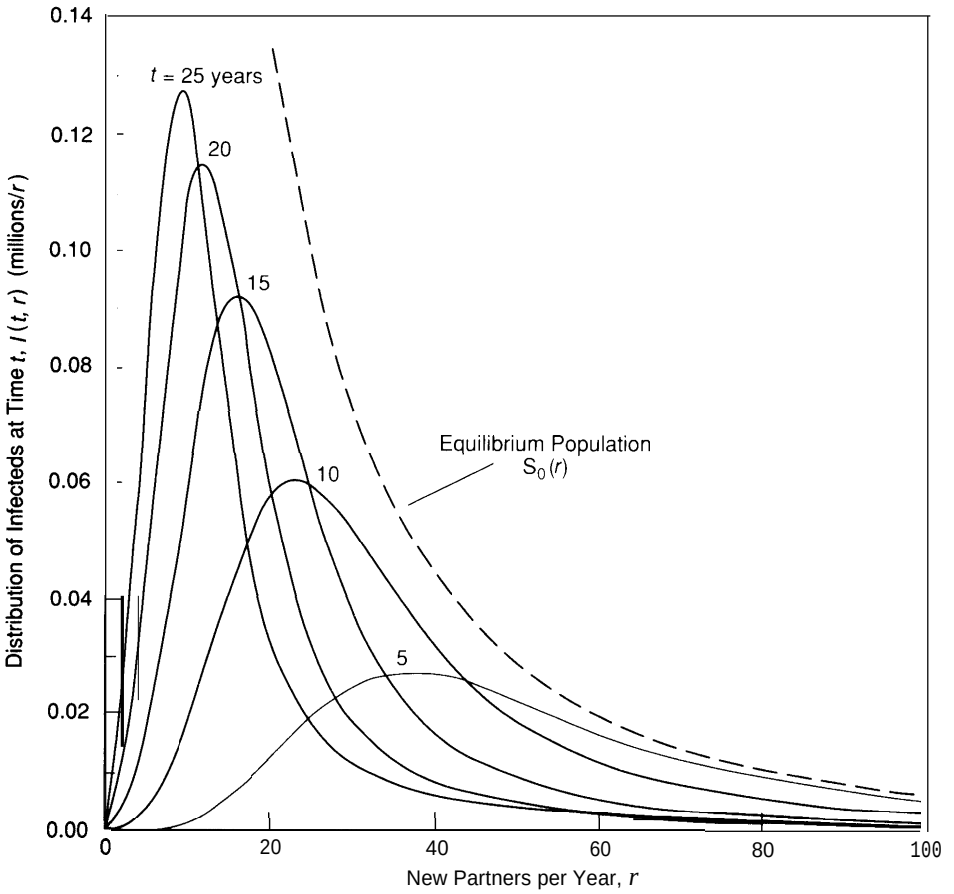
Fig. 3. The mean infectiousness $i(\tau)$ versus time since infection (solid line) used in all but the last solution presented here. The function $i(\tau)$ is an average over individuals each of whom develops AIDS at some time between 2 and 20 years since infection. The average infectiousness of each individual over the time from infection to AIDS is 0.025. The dotted line shows the pattern of Infectiousness that we postulate for a single individual. In this case the individual develops AIDS 8 years after infection. The initial peak of Infectiousness for this Indlvldual is always taken to be greater than 6 months because our numerical techniques are not yet designed to handle sharper peaks.





BASELINE SOLUTION

Fig. 4. The time-dependent behavior of various sectors of the population predicted by the baseline calculation. Despite a slow migration of people into the total population, the high mean new-partner rate of 24 partners per year drives an epidemic that substantially depletes the total population as a large fraction become infected and then die of AIDS. The very slow progression from infection to AIDS and rapid death from AIDS produces a delay between start of infection and the AIDS epidemic. Also, at all times many fewer people have AIDS than are infected.



SATURATION WAVE IN BASELINE SOLUTION

Fig. 5. Distributions of the number infected over number of new partners per year at times t = 5, 10, 40 years during the baseline calculation. The dotted line shows the distribution of the total population in the absence of HIV. As time progresses, a wave of infection moves from high-risk to low-risk groups. Essentiality all members of high-risk groups become infected, and the populations of those groups decrease to very low levels as everyone develops AIDS and dies. As the wave moves progressively through lower-risk groups, an ever smaller fraction of those groups becomes infected.

factiousness per contact since time from infection $i(\tau)$.

We present first a "baseline" solution. The acceptance function $f(r, s)$ and the infectiousness per contact $i(\tau)$ for this solution arc described in Figs. 2 and 3, respectively. The acceptance function in Fig. 2 is an inverse quartic function of r and s, which describes the probability that a person with risk behavior r. chooses a partner with risk behaviors:

$$f(r,s) = \left[1 + \frac{(r-s)^4}{\epsilon(r+r_m)^4}\right]^{-1},$$

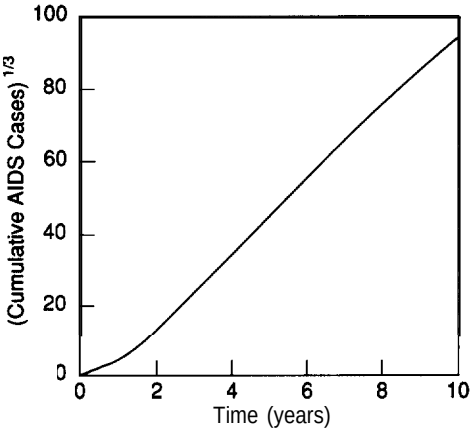
where $\epsilon = 0.01$ and $r_m = 10$ partners per year. The figure shows $f(r, s)$ versus s for three different values of r. As r increases, the width of the acceptance function increases. In rough terms, this function describes a biased mixing pattern in which a person with risk r chooses most of his or her partners from a group that ranges in risk behavior from $\frac{1}{2}r$ to $2r$.

Figure 3 is a plot of i(T), the mean infectiousness per partnership versus time since infection. The mean infectiousness is an average over the infectiousness of many individuals each of whom develops AIDS at different times (determined by $\gamma(\tau)$) since the time of infection. Figure 3 also shows the infectiousness curve for an individual who develops AIDS 8 years after infection. The infectiousness for this individual is assumed to have an initial peak, a latency period of about four years, and finally a steady rise. The average infectiousness for each individual is assumed to be 0.025. The initial peak is about 6 months wide, probably too wide to be realistic, but our numerical code does not yet have the capability of resolving a burst that is only a few weeks in duration. Nevertheless, the wider shape that we have used serves the purpose of illustrating what the impact of an initial peak of infectiousness can be.

The infected population at $t = 0$ contains 1000 individuals distributed as a narrow Gaussian function of r centered at 175 partners per year and distributed linearly in τ. Although here we assume that the epidemic starts among the highest-risk groups, this choice does not have a major impact on the numerical results. In particular, if the infecteds at t = O are centered at the mean, the epidemic follows a similar course but starts about 2 years later. If the infecteds at t = O are distributed over all risk groups, the saturation wave takes off sometime between O and 2 years later.

The input parameters and initial conditions just described yield our "baseline" solution. Figure 4 shows $S(t)$,$I(t)$, and $A(t)$ over a 40-year period. During that period about half of the population dies of AIDS. The number infected $I(t)$ and the number of people with AIDS at any given time $A(t)$ rise steadily for more than 10 years and then decline slightly as the epidemic reaches a steady state.

Figure 5 shows plots of the number infected versus risk behavior at times t = 5, 10, 15, 20 and 25 years. Here we see that the infection travels as a saturation wave from high- to low-risk groups. The wave takes 20 to 25 years to reach the lower-risk groups.

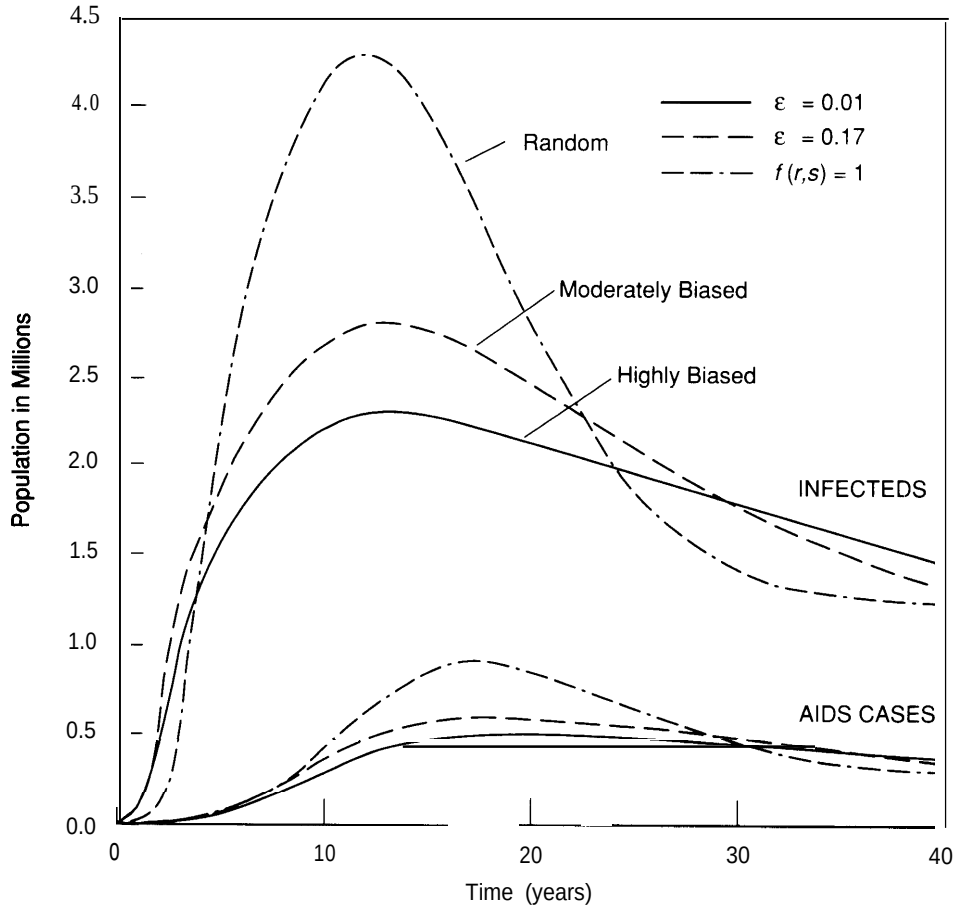


"CUBIC GROWTH" OF BASELINE SOLUTION

Fig. 6. The cube root of the cumulative number of AIDS cases as a function of time for the basellne solutlon. Although the curve is not perfectly straight, a t^3 growth in the cumulative number of AIDS cases Is a good flt to thls calculatlon between t = 1 and t = 9 years. Thus, desplte the many complexities included In the numerical model, its solutions behave quite slmllarly to the analytlc calculatlon of the main text. Note that the calculated time scales are fixed by the average value we assume for the product $c(r, s)I(\tau)$ and are therefore highly uncertain.

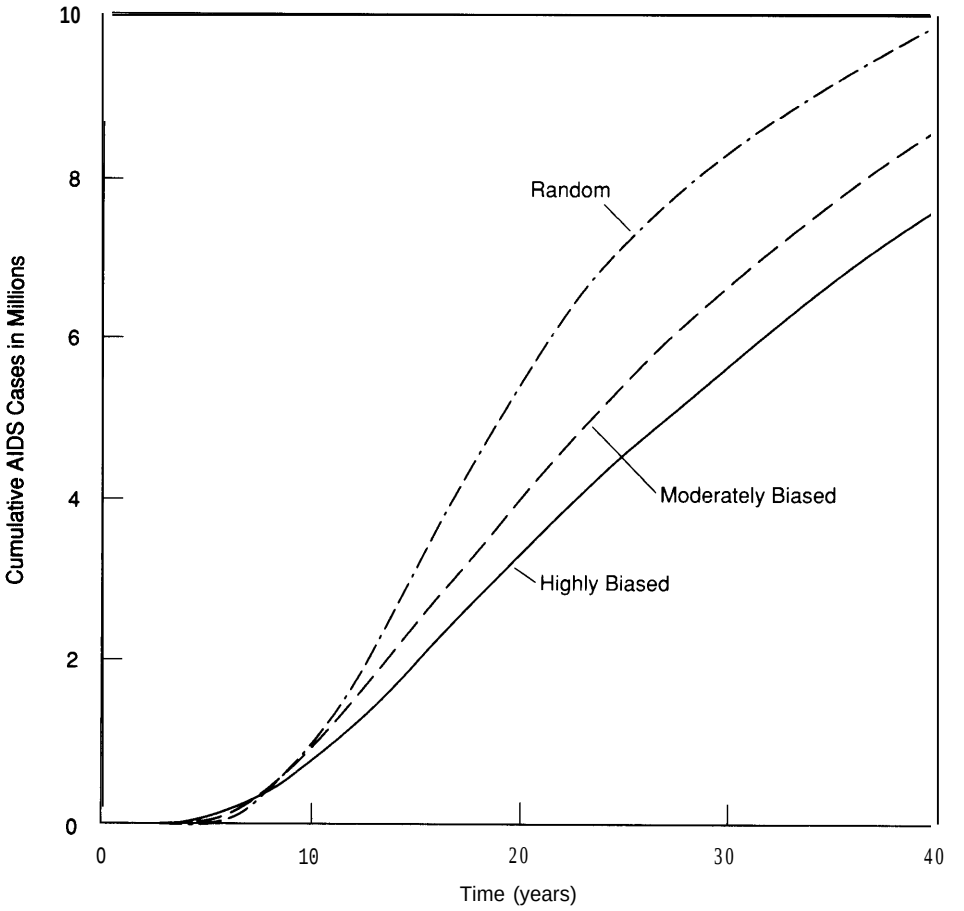
Figure 6 is a plot of the cube root of the cumulative number of AIDS cases as a function of time. The nearly straight line between 1 and 10 years shows that the calculation is not inconsistent with the observation that the number of AIDS cases grows as t^3 during the initial stages of the epidemic. The main reason that the growth is not purely cubic is the deviation of the initial profile $S_o(r)$ from a pure inverse cubic. However, the profile we chose for $S_o(r)$ fits the available partner-change-rate data much better than does Eq. 13 in the main text. We have also assumed a fairly large infectivity, which speeds up the progress of the entire epidemic. Consequently, by 10 years from the start of the saturation wave, the wave front has reached the lowest-risk populations, which, in turn, slow down the cubic growth. Although the solution just presented roughly matches the observed cubic growth of AIDS, it does not prove that the input parameters are correct but rather suggests the basic ingredients needed to produce the type of epidemic we are experiencing. A slightly different mix of input parameters yields very similar growth.

The assumption of biased mixing is the feature that sets this model apart from other models. Let's see how the epidemic changes when this assumption is relaxed. Figure 7 shows three solutions to the model that differ only in the



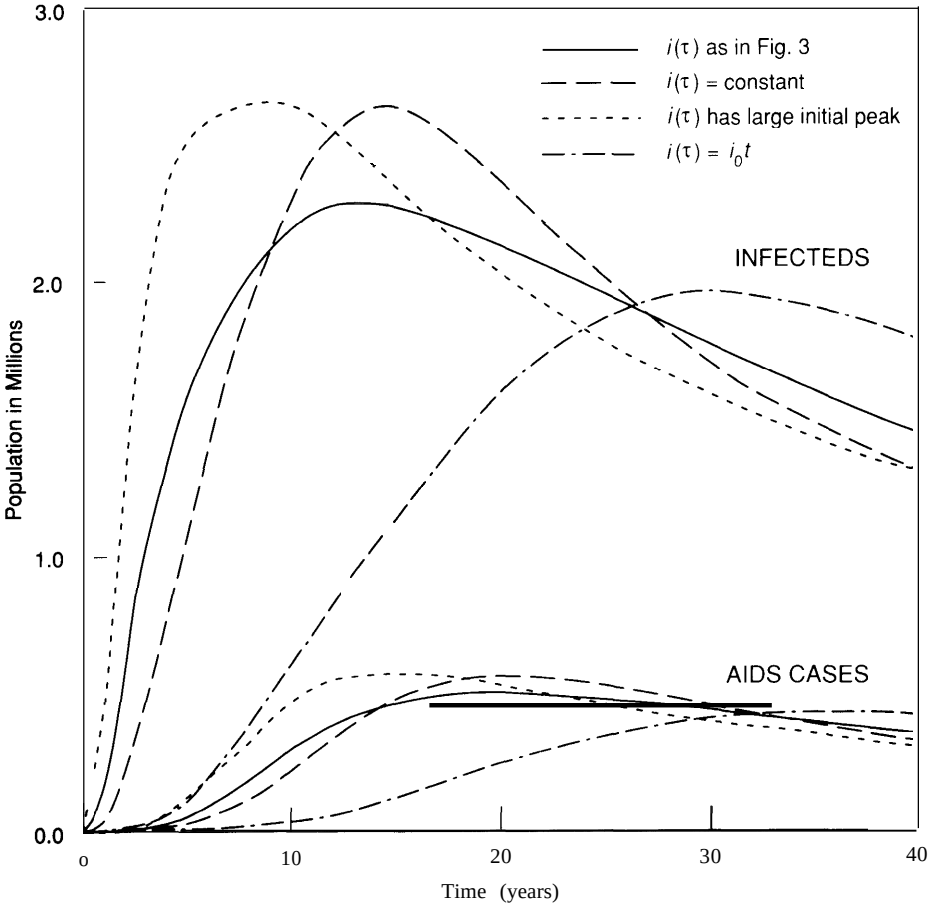
BIASED VERSUS RANDOM MIXING

Fig. 7. Time-dependent behavior of the number infected and the number of AIDS cases for various degrees of mixing among people with different risk behaviors. The baseline calculation (solid line) corresponds to the highest bias, or narrowest range of mixing. As the range of mixing widens, the epidemic changes dramatically. The growth pattern of the number infected appears to change more than that of the AIDS cases partly because of the scaie of the piot, and partly because the siow conversion to AIDS smears out the effects of the change in the number infected. More biased mixing produces a more rapid initial growth than does random mixing, but growth slows down as the infection spreads among low-risk people and the total epidemic is smaller than that produced by random mixing. When mixing is random, high-and low-risk people, pass the virus back and forth between them, so an infected person is much more likely to encounter an uninfected person until the whole population saturates.



CUMULATIVE GROWTH IN AIDS AS MIXING VARIES

Fig. 8. Cumulative AIDS cases versus time for the calculation in Fig. 7. When mixing is random, the cumulative number of AIDS cases grows exponentially until the entire population reaches saturation of infections. When the mixing is highly biased, the number grows more as a polynomial.



EFFECTS OF VARYING THE INFECTIVITY

Fig. 9. The distribution of number infected $i(\tau)$ as a function of new-partner rate at $t = 10$ years for the calculations in Fig. 7. This figure demonstrates most dramatically the effects of varying the mixing patterns. When people have a strong bias to mix with others of similar risk, few people of low risk are Infected In the early stages of the epidemic. In contrast, when partners are chosen purely on the basis of availability, people of low risk are infected early. The fact that early AIDS eases and early cases of infection wars among people with high new-partner rates is evidence for biased mixing in the U.S. population.

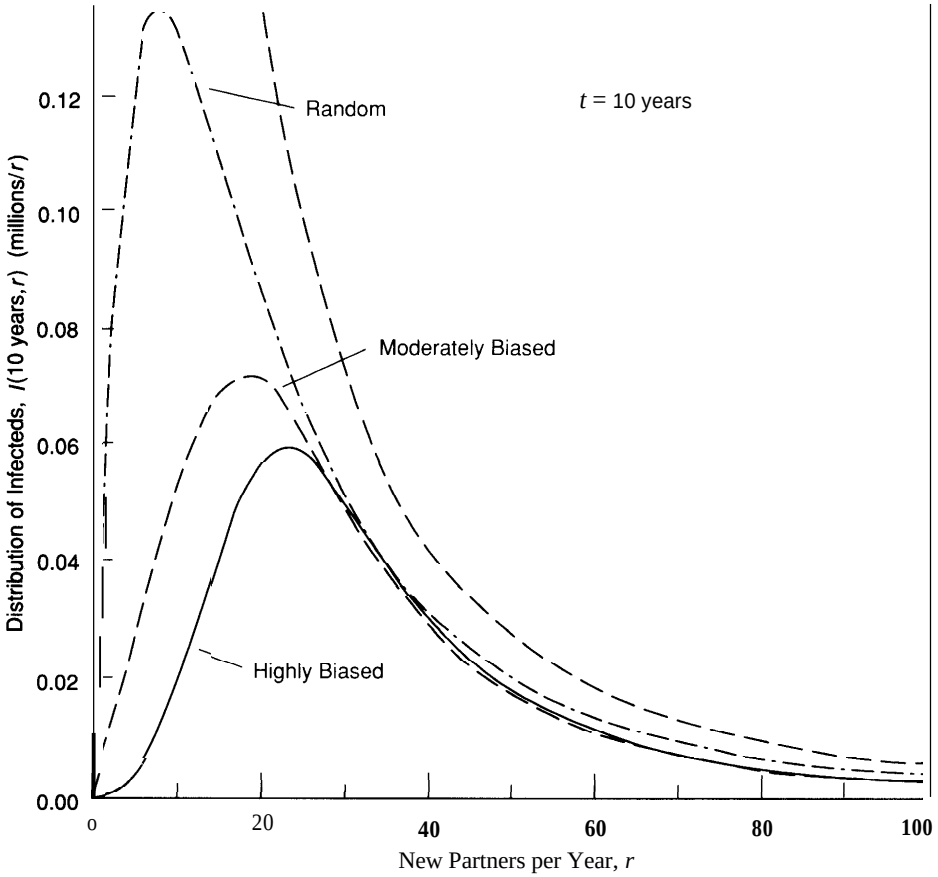
level of mixing among different risk groups. The solid lines show the baseline solution in which the mixing is strongly biased; that is, *f(r, s)* is an inverse quartic with $\epsilon = 0.01$ (see Fig. 2). The dotted lines show a solution with less bias; that is *f(r, s)* is again an inverse quartic but $\epsilon = 0.17$ so the curves of *f(r, s)* versus *s* for different values of *r* have much wider peaks than those in Fig. 2. The dashed lines show a solution with no bias; that is, $f(r, s) = 1$ corresponding to random, or homogeneous, mixing. Note that as the mixing becomes less biased, the epidemic starts off slightly later but then grows faster because the doubling time increases at a slower rate.

Figure 8 shows the cumulative number of people with AIDS as a function of time for the three types of mixing. For random mixing, the number of people with AIDS grows nearly exponentially; that is, the doubling time is nearly constant. As the mixing becomes more biased, the number of people with AIDS grows more like a low-order polynomial.

It is worth cautioning that the initial distribution of infecteds, which is arbitrary, can have a significant impact on the early growth of the epidemic, especially if the initial growth rate is low. For the random-mixing case, growth in infections is so low initially that most people getting AIDS in the first 10 years were infected at $t = 0$. Consequently, since those infected at $t = 0$ were distributed linearly with τ, the number of AIDS cases grows as a polynomial during the first 10 years, and only the number infected grows exponentially. After 10 years both the number infected and the number of AIDS cases grow exponentially. For the cases of more-biased mixing, the initial growth in number of infecteds is more rapid, so the initial distribution $I(0, \tau)$ affects the solutions for a shorter period of time. Since our initial conditions are arbitrary, rather than based on knowledge of the earliest stages of the epidemic, the solution transients just described are also arbitrary.

Figure 9 shows the number infected versus risk behavior at $t = 10$ years for each of the three mixing patterns. We see that random mixing not only produces a faster-growing epidemic but also causes the epidemic to reach the low-risk groups almost immediately. Figures 4a and 4b of the main text also illustrate that point. The solution with biased mixing shows a saturation wave of infection traveling from high-to low-risk groups, but the solution with homogeneous mixing shows no such wave. Instead, the majority of those infected are always in the low-risk groups. Since the average partner rates for the earliest AIDS cases and infected homosexuals were high compared to the mean in the general homosexual population, these numerical results support the conclusion in the main text that biased mixing has produced the cubic growth of the AIDS epidemic.

We will now examine the effects of varying the function $i(\tau)$, the infectiousness since time of infection. In the main text we used a constant value of $i(\tau)$, but we also discussed the effects of a variable infectiousness. Here we display four solutions, each of which uses a different function for $i(\tau)$ (see Fig. 10). In all cases the mean infectiousness of an individual over the course of infection is 0.025. The solid lines correspond



INFECTEDS VERSUS RISK AS MIXING VARIES

Fig. 10. Tlme-dependent behavior of the number infected and the number of AIDS cases for various assumptions about the time-dependence of Infectiousness. In these calculations we assign the same value for the average invectivlty of any individual over the course of the epidemic and vary only the distribution of infectivity with time. A burst of infectivity Just after infection causes the disease to spread very rapidly in the high-risk groups but has less effect as the disease spreads to groups with lower new-partner rates. A slowly rising infectivity several years after the initial burst sustains the epldemlc In low-risk groups. With no initial burst of infectivity, but only a slow increase from Infectlon until death, the epidemic initially spreads very slowly, but as more people approach the later stages of Infection, the epidemic gains momentum. Without control measures the epidemic may eventually affect as **many people as the other examples shown In the figure.**

to the baseline solution shown earlier; $i(\tau)$ for that solution is shown in Fig. 3. The dashed lines are the solution when $i(\tau)$ is constant. The dotted lines are the solutions when the infectivity of a person getting AIDS at 8 years has a very large initial peak, then a 4-year period during which $i(\tau) = 0$, and finally a slow increase in $i(\tau)$ up to 8 years since infection. The dash-dot lines are the solutions when $i(\tau)$ has no initial peak, but instead, a person's infectivity increases continuously between the time of infection and the time of AIDS. A large initial peak in $i(\tau)$ produces the fastest-growing epidemic, the absence of an initial peak produces the slowest-growing epidemic, and a constant value for $i(\tau)$ produces an epidemic that is closest to the baseline solution but grows a bit more slowly at first, then somewhat faster, and finally approaches a similar steady state. (Note that the vertical scale in Fig.10 is a blow up of the vertical scale in Fig. 7.) In all cases the growth is "polynomial" in that the doubling times increase continuously. Nevertheless, the shape of $i(\tau)$ has a significant impact on the course of the epidemic.

Without better data for $i(\tau)$, the future course of the present epidemic cannot be estimated. Similarly, adequate data on the mixing patterns among different risk groups is sadly lacking. If nothing else, our risk-based model points out the areas for which more data are needed. We hope that this work will help to guide the data collection and analysis efforts that are now under way. ■