The Seeding Wave

by Stirling A. Colgate and James M. Hyman

Let us assume that our risk-baaed model is a reasonable description of how AIDS has grown since the time when a member of the highest risk group was infected. In other words, we assume the infection spread as a saturation wave from the highest risk group down through lower and lower risk groups. The question remains-what happened *before* the start of the saturation wave? Did an individual from the highest risk group become infected first and start the saturation wave immediately, or did an individual from a much lower risk (and therefore much larger) group start a slow spread of infection from lower to higher groups prior to the saturation wave? We call a slow spread of infection from lowest to highest risk groups a *seeding wave.* Now, if a seeding wave *were* started, do subsequent seeding events circumvent the slow spread by leapfrogging the infection to the highest risk group, thereby reducing the number infected before the start of the saturation wave?

Here we present a model of a seeding wave consistent with our saturation-wave model of subsequent growth. In particular, the model incorporates the same distribution of risk behavior and assumptions about biased mixing used in our risk-based model. We argue that, provided these assumptions are correct, the seeding wave is a likely scenario for the early spread of HIV infections in the United States. Moreover, the model predicts that the earliest HIV infection occurred in the mid-sixties, a prediction consistent with the first recognized case of AIDS in St. Louis in 1969.

Early Growth. Suppose the first infection in the United States is initiated, say, by either a visitor with HIV or a U.S. person visiting elsewhere. Although these two cases would not be equivalent if high risk of infection is correlated with high rate of travel, we will not consider such correlations here, Rather, we assume that risk of infection can be quantified using a single variable r with its distribution $N(r)$ defined by Eq. 13 in the main article. Since the probability of a person becoming infected is proportional to r, the probability $P(r)$ that at least one individual with risk r or greater becomes infected is given by

$$P(r) \propto \int_r^\infty N(r)rdr = r^{-1}, \qquad (1s)$$

for $r \geq 1$, that is, for the high-risk end of the population defined by Eq. 13.

Hence, the smaller r is, the greater is the probability that at least one individual of risk group r becomes infected despite the lower risk per individual. Also, the most likely case is that the first infected individual was a member of the average group, the group with $r = 1$.

A Simple Numerical Model. We wish to model the progression of the infection to higher risk groups starting with an infected individual close to the average, To help understand this process, we simplify by saying that the *kth* risk group r_k varies in risk behavior by a factor of 2, that is, r varies from r& to $2r_k$. Hence, the various groups will have risk behaviors 1, 2, 4, 8... times the average. The number of individuals for $r > 1$ in the *kth* group (using Eq. 13) is

$$\int_{r_k}^{2r_k} N(r)dr = -\frac{1}{2}N_0 r^{-2}\Big|_{r_k}^{2r_k} = \frac{3}{8}N_0 r_k^{-2}. \qquad (2s)$$

Since the total population (the integral of Eq. 13 from $r = 0$ to $r = \infty$) is $\frac{3}{2}N_0$, our first group with $1 < r < 2$ is one-fourth of the total, and if we restrict ourselves to the homosexual population, one-fourth of the total is one million. Thus, the second group, with $2 < r < 4$, will have $(\frac{1}{4})(\frac{3}{8})N_0$ individuals, or $\frac{1}{16}$th of the total population, or 250,000. The third group, with $4 < r < 8$, will have $(\frac{1}{16})(\frac{3}{8})N_0$, or $\frac{3}{64}$th the total population, and so forth.

We do not believe that people exhibiting a preference for each other are likely to recognize a behavior difference much finer than a factor of 2, and hence, we use this rather crude measure of a group. We also suppose, as a reasonable but unknown example, that the fraction of the time an individual participates in risk outside his group is $F = \frac{1}{4}$. If this fraction is greater or less by a factor of 2, it will change what follows by a factor of 2, but that change is within the accuracy of these estimates.

We next ask how many individuals must be infected in group 1 before a member of group 2 is infected. There are one-fourth as many people in group 2 as in group 1 with twice the risk behavior; that is, the number of each group decreases as $1/r_k^2$ (from Eq. 2s), and they have contacts with other groups only one-fourth of the time. This fraction of out-of-group mixing will be distributed between both higher and lower risk groups.

Let us assume that the fraction is evenly divided between the higher and lower groups and, because of the same bias that leads to group preference, is primarily in the adjacent groups. Crudely then, *F can be* considered to be a diffusion coefficient. A bias towards only adjacent out-of-group mixing prevents the infection from leapfrogging to much higher groups and circumventing the slow seeding-wave progress.

The seeding wave progresses from group k to the next higher group $k + 1$ when one member of the next higher group is infected. If I_k is growing exponentially, $I_k = e^{(1-F)\alpha r t}$, then the cumulative probability of infecting one member of group $k + 1$, starting at the time when one member of group k is infected, is

$$\begin{aligned} I_{k+1} &= \int_0^t \frac{F}{2}(1-F)(\alpha r_k) I_k \, dt \\ &= \frac{F}{2} e^{(1-F)\alpha r_k t}\Big|_0^t \\ &= \frac{F}{2}(I_k - 1), \end{aligned} \tag{3s}$$

where the factor $\frac{F}{2}$ is needed because only one half of the out-of-group mixing pertains to the higher risk groups. The remaining half augments the growth rate of the next lower risk group.

Since $I_{k+1} = 1$ when the seeding wave progresses by one group, I_k at this transition becomes equal to $\frac{2}{F+1}$. Therefore, in our example (for which $F = \frac{1}{4}$), *nine* members of a group must become infected before a member of the next higher group becomes infected. The time for this to occur will be

$$t_k = \frac{\ln(\frac{2}{F} + 1)}{(1-F)\alpha r_k}. \tag{4s}$$

Thus, the speed of the seeding wave is *dk/dt = 1/tk. The* remaining time for the seeding wave to go from group k to the highest risk group at $k = m$ *is*

$$t_{km} = \sum_k^m t_k, \tag{5s}$$

or

$$\begin{aligned} t_{km} &= \frac{\ln(\frac{2}{F} + 1)}{(1-F)\alpha} \sum_k^m \frac{1}{r_k} \\ &\simeq 2\frac{\ln(\frac{2}{F} + 1)}{(1-F)\alpha}, \end{aligned} \tag{6s}$$

for $m >> k$. That is, the sum $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \cdots \simeq 2$ after even just a few terms. Thus, for most of the groups with $k < m$, the remaining time needed for the seeding wave to move through essentially all groups (that is, all but the few of highest risk) is just double the time to infect the adjacent next higher group. Now, each of these seeded groups is growing exponentially so that, as the time increases from, say, t_k to $2t_k$, the number infected in the *kth* group increases from I_k to I_k^2. Thus, the number of individuals infected in each group at the time the seeding wave ends, $t = t_m$, will be the square of the number infected when the next higher adjacent group is seeded with one individual, or $(\frac{2}{F} + 1)^2$. That is,

$$\begin{aligned} I_k(t_m) &= \int_0^{t_m} e^{(1-F)\alpha t} dt \\ &\simeq (\frac{2}{F} + 1)^2 \quad \text{for } k \ll m. \end{aligned} \tag{7.}$$

Since the seeding wave progresses through m groups and each group has one-fourth the members of the next lower risk group, $m = \ln \frac{N_0}{4} / \ln 4 \simeq 11$, for a total host size of 4 million, or $\frac{3}{2}N_0$. Thus, the maximum number likely infected at the start of the seeding wave is $m(\frac{2}{F} + 1)^2 \simeq 860$. Of course, the out-of-group mixing fraction F is only poorly estimated, and a factor of two larger or smaller value for F implies a range of 270 to 3000 infected before the start of the saturation wave. Although these estimates cover a wide variation, they are upper bounds on the number infected before the start of the saturation wave. As mentioned above, leapfrogging would circumvent the seeding wave and reduce the number infected prior to the start of the saturation wave. Moreover, these upper bounds are not inconsistent with the prediction in the main text that the size of the infected cohort before the start of the saturation wave, namely I_0 in Eq. 24, is small.

This very simple description of the initial spread of infection opens up a number of questions. (1) What is the likely time when the first individual

was infected and, hence, later became the likely first case of AIDS? (2) Is the predicted risk behavior of the early cases of AIDS, inclusive of the seeding wave, consistent with the high mean risk behavior of the early AIDS cases observed by the Centers for Disease Control (CDC)? (3) What is the probability that the whole process of group-to-group progression is circumvented by one high-risk individual becoming infected early in the seeding process? (4) Is the seeding process consistent with our perception that all major demographic groups participated in a near simultaneous start, that is, synchronization of the saturation wave?

Infection Time. We would like to associate a real time with the time step t_k of Eq. 4s and then take the sum $\sum_1^m t_k$ as the total, or maximum likely, time of the seeding wave. This then becomes the ***maximum*** time prior to 1979.2 that the first person in the United States was likely to have been infected.

In the seeding-wave process, the growth rate of any given group is $(1 - F)\alpha r_k$, where the factor $(1 - F)$ recognizes that out-of-group mixing is not balanced by equal and opposite in-group mixing. We now use the current growth rate of the group at the front of the saturation wave to calibrate $(1-F)\alpha$. In this way, we derive a very rough estimate for the maximum time of the seeding wave.

Figures 2 and 3 of the main article indicate that, at the time 1988.2, the homosexual fraction was approximately 65 per cent of our estimated one million infected, which is 650,000 infected, or 1/6 th of our estimate of the total number of active homosexual population of 4 million. This estimate places the presently infected population partly in group 1 with all higher groups near saturation. The total population already infected in the higher risk groups is $\frac{4}{3}N_2$, or roughly 330,000 (Eq. 2s). Thus, about the same number must be infected in group 1 so that the total is 650,000.

It has required 9 years for the seeded fraction of 81 individuals in group 1 to grow to 320,000, which gives a growth rate of $(1 - F)\alpha = \frac{1}{9} \ln \frac{320000}{81} = 0.92$ per year or a doubling time of 0.75 years. Thus, the apparent growth rate for the total epidemic, which must be averaged over both group 1 and all higher risk groups—groups that, by now, are almost saturated, gives a doubling time that is roughly twice as large, or 1.5 years. This doubling time is to be compared to the present doubling time for infection predicted by our saturation-wave model, which, at t + 2 = 9 years, is $(\frac{1}{I}dI/dt)^{-1} \ln 2 = 0.69t/2 = 3$ years. The three-year doubling time corresponds to a two-year doubling time for AIDS, in agreement with present CDC estimates of 1.75 years. Thus, our saturation-wave model is consistent with the CDC data but inconsistent with the simple seeding-wave growth by a factor of two. One source of discrepancy is our incomplete treatment of the effects of out-of-group mixing. We therefore estimate that the growth rate in group 1 is bounded by a doubling time of 0.75 to 1.5 years.

In Eq. 2s we have neglected group O $(O < r < 1)$ with 3.3 million individuals. The first individual infected is equally likely to be in group O or 1 because the average value of $N(r)r$ ***is*** approximately the same for both groups. We neglected group 0 to simplify the seeding-wave calculation, but since our estimates for the doubling time are too short, we must now recognize that the initial infected individual most likely had a lower mean risk than group 1 and that the mean growth rate is between the growth rate of two groups. As a rough approximation, let us say that the mean growth rate is lower by a factor of $1/\sqrt{2}$. Then the doubling time of the combined group average will be $0.75\sqrt{2}$ to $1.5\sqrt{2}$ years, or 1.1 to 2.2 years. This then becomes a rough estimate of the doubling time of the seeding wave.

First Infection. The date for the beginning of the saturation wave or power-law (t^2) growth of infection was 1979.2 (Eq. 24). But the seeding-wave model suggests that the first infection in the United States may have occurred $\ln((\frac{z}{F} + 1)^2)/\ln 2 \simeq 6$ doubling times earlier, or 7 to 14 years earlier. The date of the first infection thus may fall somewhere between 1972 to 1965, earlier than has previously been estimated. Thus, the singular case of a teenage boy in St. Louis who has now been identified as having died of AIDS in 1969 is consistent with our seeding-wave picture if he was infected up to five years before developing AIDS. The existence of this case of AIDS in 1969 implies a slow growth of the number infected before the start of the saturation wave.

Mean Risk Behavior. We wish to confirm that our model of the seeding wave, which starts in relatively low-risk groups, is consistent with the CDC observation that most early cases of AIDS were among high-risk individuals. The mean risk behavior of those developing AIDS can be calculated using a convolution integral similar in structure to Eq. 26, but one emphasizing risk rather than time since infection. However, here we are really interested in risk behavior versus time and the absolute number of cases of AIDS, because it was the occurrence in 1981 of roughly 50 AIDS cases in a relatively short period of time (approximately 6 months) that caused the recognition of an epidemic.

We next must define high- and low-risk behavior in terms of our seeding-wave model. The new-partner rate of the homosexual population in London SDT clinics (Fig. 5 in the main text) has a mean of roughly 24 partners per year.

We associate this new-partner rate with group 1 of the seeding-wave model. Group 2 would then have a mean rate of 48 new partners per year-well within the CDC definition of extremely high risk behavior. Thus, moderate or low risk behavior is restricted to groups 1 and 0 with doubling times of 1.1 to 2.2 years and 0.8 to 1.6 years, respectively. By 1979 these two groups would each have infected $(\frac{2}{F}+1)^2 \simeq 81$ individuals. Two years later, in 1981, the combined groups would be producing AIDS cases at a rate of 6 per cent per year, that is, 0,06 x 2 x 81, or 10 cases per year. The total cases for 1981 was several hundred, so these 10 additional moderate-risk cases would, by comparison, be negligible. Thus, we believe that the seeding-wave model is consistent with the CDC observation that high-risk behavior was strongly correlated with the AIDS cases at the start of the epidemic.

Bypassing the Seeding Wave. Of course, this slow growth for 7 to 14 years in group 1 could have been bypassed by one member of any group with $r \geq 2$ becoming infected at the beginning. The probability of this happening per infection in group $k-1$ is, for each group, proportional to $N_k r_k \propto F/(2r_k)$ per group, discounting mixing biases. Therefore, the probability that at least one member of higher risk becomes infected, exclusive of the seeding wave, becomes

$$P = \frac{r}{2}\sum_{2}^{-m.}\frac{1}{\cdot\cdot} \cong \frac{r}{2}\cdot \qquad (8s)$$

That is, when one member of group 2 becomes infected, it is equally likely that a member of any higher risk group will become infected, and then the remaining time to the start of the saturation wave becomes negligibly small. This effect would reduce the time for the start of the saturation wave by a factor of 1/2 or less, or just the time to infect one member of group 2, which is within the error of our estimates. On the other hand, if we wish to preserve this factor of 2, we must require that F is a function of r_k/r_{k+n}. A bias function as weak as $F \rightarrow F/\ln(r_{k+n}/r_k)$ guarantees that the roughly 100 out-of-group infections likely to occur during the course of the seeding wave have a small probability of being in the highest risk group ($k+1 \leq m$). Otherwise, infection will leapfrog to reach saturation in one-half the time.

The seeding time would likewise be shorter if the infectious source (in another country, for instance) grew rapidly enough to cause many infections in group 1 and, hence, at least one infection in higher risk groups. There is also the possibility that the infection started and died out several times in groups O and 1 before starting the seeding wave. This possibility is equivalent to saying that the net reproductive rate of the disease is very close to unity in these low-risk groups, that is, that a given infected individual infects only one other in the mean time of 10 years to AIDS and death. Because of the arguments in the main text concerning the probability of infection per sexual contact and the equivalence of new-partner rate and contact frequency, we believe the net reproductive rate in the homosexual population was large, and thus the seeding wave started with the first infection.

Synchronization of Risk Populations. We ask if either the slow seeding wave or the singular high-risk initial case of infection makes any difference to the saturation-wave model. For sexual preference and race (Figs. 3 and 4 in the main text) as well as regional and age populations (not shown), the cube root of the cumulative number of cases is nearly linear for $t \geq 1982.5$. These curves extrapolate to zero at approximately the same time with a maximum delay of half a year. This result means that all these subpopulations had to be seeded with at least one high-risk member infected within this time interval. The number infected after half a year (using Eq. 24) becomes roughly 2000 or 3000, all within high-risk categories and with or without the seeding wave of initial infection. We then ask what the probability is that any population selected does not have one member within the 2000 to 3000 initially infected high-risk groups. This depends upon the social isolation, but for a subdivision that creates 10 or more categories, no one population is likely to have less than 100 to 150 members in its high-risk group. Thus, isolation would have had to been very strong, such that none were infected. The observed synchronization of the subpopulations seems reasonable and is independent of whether a seeding wave or single high-risk infection started the epidemic.

In summary, we have described a plausible process by which the initial infection of HIV spread in the various risk populations of the United States. Initially, an average individual was infected from sources unknown, but the infection then grew in a peer group until the number infected and the probability of out-of-group mixing caused the infection to jump to a higher risk group. In this fashion, a seeding wave of infection steadily climbed to the highest-risk individuals. The rapid growth among these highest-risk individuals caused all of them to be rapidly infected, resulting in the start of saturation-wave growth for the whole population. The total number infected in the initial seeding wave is strongly dependent upon the out-of-group mixing fraction, but reasonable estimates indicate that the number infected by the seeding wave would be small enough, less than several thousand, to leave the later saturation-wave growth intact. The earliest known case of AIDS in the U.S. in 1969 is consistent with this picture. ■