

NONRANDOM MIXING MODELS OF HIV TRANSMISSION*

Edward H. Kaplan, Peter C. Cramton, and A. David Paltiel
Yale School of Organization and Management
Box 1A, New Haven, CT 06520, USA

Abstract

Models of HIV transmission and the AIDS epidemic generally assume random mixing among those infected with HIV and those who are not. For sexually transmitted HIV, this implies that individuals select sex partners without regard to attributes such as familiarity, attractiveness, or risk of infection. This paper formulates a model for examining the impact of nonrandom mixing on HIV transmission. We present threshold conditions that determine when HIV epidemics can occur within the framework of this model. Nonrandom mixing is introduced by assuming that sexually active individuals select sex partners to minimize the risk of infection. In addition to variability in risky sex rates, some versions of our model allow for error (or noise) in information exchanged between prospective partners. We investigate several models including random partner selection (or proportionate mixing), segregation of the population by risky sex rates, a probabilistic combination of segregation and random selection induced by imperfect information (or preferred mixing), and a model of costly search with perfect information. We develop examples which show that nonrandom mixing can lead to epidemics that are more severe or less severe than random mixing. For reasonable parameter choices describing the AIDS epidemic, however, the results suggest that random mixing models overstate the number of HIV infections that will occur.

1. Introduction

The AIDS epidemic has now claimed almost 55,000 lives in the United States, while the U.S. Centers for Disease Control report roughly 94,000 confirmed cases of AIDS as of April 1989 (CDC, 1989). Though it is now established that risky sex (primarily unprotected receptive anal intercourse with an infected insertive partner) and the sharing of drug injection equipment (e.g. needles) are the major modes of transmitting human immunodeficiency virus (or HIV, the causal virus of AIDS; see Curran et al, 1988), surprisingly little is known about the basic quantities which are driving the AIDS epidemic. Critical facts which have not been firmly established but are thought to be important include the distribution of risky behavior rates across the population (Anderson et al, 1986; Hay and Anderson, 1987), the manner in which different subpopulations mix (Hyman and Stanley, 1988; Abramson and Rothschild, 1988; Hethcote and Van Ark 1987), the duration of time for which such behaviors are practiced (which is closely linked to the AIDS incubation time; Anderson et al, 1986; Hay and Anderson, 1987; Kaplan, 1989a; Lui, Darrow and Rutherford, 1988; Brookmeyer and Gail, 1988), and the likelihood of becoming infected given exposure to the HIV virus (Grant, Wiley and Winkelstein, 1987), perhaps adjusting for mode of exposure and variability in infectivity over time and across individuals (Anderson, 1988; Hay and Anderson, 1988; Wiley, Herschkorn and Padian, 1989).

In two important papers, Anderson et al (1986) and May and Anderson (1987) proposed a simple model of the AIDS epidemic among gay men. Their model incorporates a distribution of risky sex rates across the sexually active gay population. Though their model thus allows for heterogeneity in risky sex rates, they assume that sex partners are selected at random from the population at large. Indeed, the assumption of random partner selection is central to most of the sexually transmitted disease and AIDS models developed to this point (see Anderson et al, 1986; May and Anderson, 1987; Anderson, 1988; Hethcote and Yorke, 1984; Isham, 1988; and Kaplan, 1989a and 1989b for examples). Some researchers have considered departures from random partner selection (Abramson and Rothschild, 1988; Blythe and Castillo-Chavez, 1989; Hethcote and Van Ark 1987;

* In Mathematical and Statistical Approaches to AIDS Epidemiology, C. Castillo-Chavez (ed.), Lecture Notes

Hyman and Stanley, 1988; Jacques et al, 1988; Koopman et al, 1988; Nold, 1980). Clearly, random partner selection is not behaviorally realistic, in that attributes such as familiarity, attractiveness and risk of infection are ignored. One might think that random partner selection would produce an epidemic that is worse than what would occur with nonrandom mixing, since random selection would eventually lead to a total mix of the population. This intuition, unfortunately, can be false: random partner selection need not overestimate the size of an epidemic.

In this paper, we consider partner selection alternatives to random mixing. The mixing mechanisms we propose are motivated by behavioral assumptions and represent a broad spectrum of matching methods. Our intention is to determine if the assumption of random mixing is robust over a variety of alternative mixing schemes. We feel this is an important subject to study, given that many policy models of AIDS rely on random mixing formulations (Kaplan 1989a, 1989b; Kaplan and Abramson, 1989). If the results of AIDS models are highly dependent upon random mixing assumptions, then public health strategies which serve to contain an epidemic driven by random mixing might not contain the real epidemic (should random mixing prove false). However, if random mixing models tend to overstate the size of an epidemic, then policies which contain random mixing epidemics should prove robust to more complex mixing behaviors.

Our paper is organized as follows. Section 2 presents our assumptions underlying HIV transmission. Section 3 presents general conditions under which an HIV (and hence AIDS) epidemic can occur in the population under the assumptions of Section 2. The behavioral principle of risk minimizing partner selection is discussed in Section 4. This principle is applied to derive epidemic models corresponding to a variety of partner selection schemes: random partner selection (or proportionate mixing) owing to no information; segregation by risky sex rate owing to perfect information; Bernoulli noise (or preferred mixing) owing to imperfect information; and costly partner search with perfect information. The epidemiological implications of these mixing functions are studied in Section 5. Our concluding remarks appear in Section 6.

2. Assumptions and Model Formulation

The model to be described is similar to the model considered by Hethcote and Van Ark (1987). At time t , the population in question contains $n(t)$ gay men. This population can be divided into m subpopulations, with $n_i(t)$ men in subpopulation i , $\sum_{i=1}^m n_i(t) = n(t)$. Immigration to subpopulation i occurs at the constant rate of N_i men per year, while the per capita exit rate in all subpopulations (due to death unrelated to HIV or AIDS, or due to emigration) equals μ per man per year. In the absence of HIV infection, then, $n_i(t)$ will approach the steady state value of $n_i = N_i/\mu$.

We assume that HIV infection has been introduced to this population; such infection is assumed to spread via risky sexual contacts. Men in subpopulation i have risky sex with an average of c_i partners per year (c_i is called a risky sex rate). In any risky sexual relationship between an infected member of subgroup i and an uninfected member of subgroup j , the uninfected man becomes infected with probability β_{ij} . The probabilities of HIV transmission (the β_{ij} 's) are assumed to apply per partnership as opposed to per sexual act, an assumption which is gaining empirical support (May and Anderson, 1988; Wiley, Herschkorn and Padian, 1989). We assume

that infected men remain sexually active and infectious for an average of D years; the quantity D depends upon both the remaining active sex life from the time of infection, and the AIDS incubation time (see Kaplan, 1989a for a detailed discussion).

When a member of subpopulation i selects a sex partner at time t , the partner selected is a member of subpopulation j with probability $q_{ij}(t)$, $\sum_{j=1}^m q_{ij}(t) = 1$, $i = 1, 2, \dots, m$. The probabilities $q_{ij}(t)$ are referred to as the mixing probabilities; they need not reflect random selection among the subpopulations. Unless stated otherwise, we assume that there exists a direct or indirect path of positive mixing probabilities connecting any two subgroups (i.e., all subgroups communicate). This implies that epidemics will either develop in all subgroups, or in no subgroup. Cases where all subgroups do not communicate may be partitioned and analyzed separately. We also assume symmetry in sexual activity: when a man from subpopulation i has risky sex with a man from subpopulation j ; the reverse also occurs. This symmetry implies that the mixing probabilities must satisfy the following law of conservation of sex:

$$n_i(t)c_i q_{ij}(t) = n_j(t)c_j q_{ji}(t) \quad i = 1, 2, \dots, m; j = 1, 2, \dots, m, \quad (1)$$

$$t \geq 0$$

Equation (1) simply states that the total rate at which men from subpopulation i have sex with men from subpopulation j balances the total rate at which men from subpopulation j have sex with men from subpopulation i . Any system of mixing probabilities proposed must obey equation (1).

3. Threshold Conditions for HIV Epidemics

There are several basic questions one may ask of epidemic models such as the one just formulated, but the most basic query is this: under what condition will the number of infected men increase over time following the introduction of HIV infection into the population? To answer this question, assume that the population has reached steady state in the absence of HIV infection (so that $n_i = N_i/\mu$), and that the mixing probabilities $q_{ij}(t)$ have stabilized at q_{ij}^* .

Let $Y(t)$ denote the vector of infected men in the population at time t following the introduction of HIV infection;

$$Y(t) = (Y_1(t), Y_2(t), \dots, Y_m(t))^T. \quad (2)$$

We define the nonnegative matrix A to contain the elements

$$a_{ij} = c_i b_{ij} q_{ij}^* D \quad i = 1, 2, \dots, m, \quad (3)$$

where q_{ij}^* denotes the mixing probabilities corresponding to the disease-free steady state. Let I be the identity matrix, and $o(y)$ be a function such that $\lim_{y \rightarrow 0} \|o(y)\|/\|y\| = 0$. With this notation, we assume that

$$\frac{dY(t)}{dt} = \frac{1}{D}(A^T - I)Y(t) + o(Y(t)) \quad (4)$$

early in the epidemic. For time periods close to the origin, this model is a close approximation to many of the AIDS models referred to earlier.

As a first order differential system, it is easy to show that the behavior of $Y(t)$ is governed by the largest eigenvalue of A (which equals the largest eigenvalue of A^T , of course). Let $\rho(A)$ denote the largest eigenvalue of A (which is guaranteed to be nonnegative, as A is a nonnegative matrix; see Strang, 1980). Standard results (e.g. Strang, 1980) then dictate that early in the epidemic, $Y(t)$ will increase exponentially if and only if $\rho(A) > 1$; if $\rho(A) \leq 1$, $Y(t)$ remains stable or decreases following an initial transient. A derivation of this threshold via consideration of endemic equilibria was presented by Hethcote and Van Ark (1987) for their model; the rough arguments presented here generalize the approach used by Anderson et al (1986) in their study of random sexual mixing.

To determine when epidemics can occur requires computing eigenvalues under the approach formulated above. However, a much simpler approach which avoids the explicit computation of $\rho(A)$ can be applied providing $I - A$ is nonsingular. Imagine introducing the infection to an infinite population of susceptibles. Let T_i be the number of ultimate infections (direct and indirect) accounted for by a newly infected type i man, assuming that all risky sex involving infected men occurs with uninfected men (owing to the infinite population of susceptibles). If $T_i < \infty$ for every i , then the infection dies out in the infinite population of susceptibles, implying that no epidemic will materialize in the “real” world. As we will now show, determining the values of T_i is straightforward as long as $I - A$ is nonsingular.

An infected type i man will remain sexually active on average for time D by assumption. The total number of risky sex partnerships involving this man thus averages $c_i D$. With probability q_{ij}^* (as the total populations of the various subgroups remain in proportion to the values of N_i), the infected type i man has a risky sex partnership with an uninfected type j man. The mean number of ultimate infections generated from such a partnership equals 0 with probability $1 - \beta_{ij}$ (that is, the infected type i man did not infect the susceptible type j man), and $1 + T_j$ with probability β_{ij} (for if the infection is passed, then the newly infected man counts for one infection, and that same newly infected man will now generate T_j subsequent infections by the definition of T_j). Thus, the T_i 's satisfy the following system of equations:

$$T_i = c_i D \sum_{j=1}^m q_{ij}^* \beta_{ij} (1 + T_j), \quad i = 1, 2, \dots, m. \quad (5)$$

Let $T = (T_1, T_2, \dots, T_m)^T$, and $1 = (1, 1, \dots, 1)^T$. Equation (5) can then be rewritten as the linear system

$$T = A (1 + T) \quad (6)$$

with solution

$$T = (I - A)^{-1}A \mathbf{1}, \quad (7)$$

assuming $I-A$ is nonsingular.

Standard results from linear algebra may again be invoked: as A is a nonnegative matrix, T is finite and nonnegative if and only if $(I-A)^{-1}$ is finite and nonnegative, which in turn occurs if and only if $\rho(A) < 1$ (Strang, 1980). Thus, if equation (7) results in $0 \leq T_i < \infty$ for every i , then $\rho(A) < 1$, indicating that an epidemic cannot occur. Alternatively, if $T_i < 0$ for some i , the previous discussion implies that $\rho(A) > 1$, signalling an epidemic. Of course, only nonnegative values of T_i have physical meaning (for one cannot generate negative infections); thus negative values of T_i should be recognized as characterizing an infinite number of infectious transmissions.

As such, a second approach to determining whether or not an epidemic will occur for a given A matrix, providing $I-A$ is nonsingular, is to evaluate equation (7) and ask whether the resulting T_i 's are all nonnegative. If the answer is yes, then an epidemic cannot occur (as $\rho(A) < 1$); if the answer is no, an epidemic will occur (as $\rho(A) > 1$). Solving equation (7) (or equivalently, the system of equation (5)) can prove much easier than directly evaluating $\rho(A)$, especially if the objective is to obtain closed-form results corresponding to epidemic thresholds for specific mixing functions, as will become clear in Section 5.

4. Behavioral Foundations for the Mixing Probabilities

In this section, we motivate a number of alternative assumptions about partner selection and derive the consequences in terms of the mixing probabilities q_{ij} (we suppress the dependence on time in this section, as our arguments will apply to $q_{ij}(t)$ at all points in time). Our purpose is not so much to realistically model the partner selection process as it is to admit a variety of alternative assumptions which encompass a wide range of mixing possibilities. In later sections, these alternatives are explored to evaluate how sensitive results under the random mixing assumption are to alternative specifications of mixing behavior.

Two important questions motivate our models of partner selection: 1) how much does an individual know about a partner's risky sex rate, and 2) how costly (in time or effort, say) is the search for partners? In the following, we assume that each man is concerned with the risk associated with his sexual contacts. In particular, we assume that a man selects partners to minimize his perceived risk of infection, given the information he has about partners, his pool of available partners, and the rate at which he engages in risky sex. Our alternative models differ in the assumptions they make about the quality of the information individuals have about their partners, and the magnitude of search costs for partners. Given alternative information and cost structures, our models will force mixing among men with different risky sex rates. Such induced mixing can be viewed as an approximation of mixing due to perfect information regarding other attributes such as attractiveness and familiarity.

In all of our models where information plays a role, we assume that the only information signal (perfect or imperfect) is the risky sex rate. Thus, men evaluate the riskiness of a prospective partner in terms of his risky sex rate, the higher the risky sex rate, the higher the perceived risk. Important extensions not considered in this paper

include variation in perceived risk due to variability in risky sexual practice (e.g., cognition of the different values of β_{ij}), and variation in perceived risk due to the duration of time for which prospective partners have been sexually active (e.g., a ten-year versus five-year history of sexual activity at the same annual risky sex rate).

In the remainder of this section, we formulate four alternative models of partner selection: (1) random mixing, which is motivated from poorly informed individuals or extremely costly search for partners, (2) segregation, which results from perfect information regarding risky sex rates and negligible search costs, (3) Bernoulli noise, which stems from imperfect information about partners and negligible search costs, and (4) costly search, which is derived from perfect information of risky sex rates, but high searching costs. Each of these models results in specific functions for the mixing probabilities q_{ij} . The epidemiological implications of these models will be examined in Section 5.

4.1. Random Mixing (or Proportionate Mixing)

Random mixing (or proportionate mixing; see Anderson et al, 1986; Hethcote and Van Ark, 1987) across risky sex rates in the population will occur when individuals have no information about prospective partners' risky sex rates, or when search costs are enormous. In either case, a man is unable to influence his perceived risk through careful partner selection; in the case of no information, all partners appear equally risky, while in the case of enormous search costs, a man will choose the first person available for risky sex since it is too costly to search out more desirable partners. Random mixing might also arise if an individual's preferences are such that factors uncorrelated with perceived risk dominate the partner selection decision. Any of these assumptions (no information, enormous search costs, or the dominance of other factors uncorrelated with risk) is sufficient for random mixing to occur, so long as we make the additional assumption that the pool of available partners for members of each subpopulation has the same distribution of risky sex rates as the total population. This assumption would be violated, for example, if men, for social reasons independent of risk minimization, segregate into subpopulations correlated with risky sex rates.

A major virtue of the random mixing model is its simplicity. With random mixing, the pairing mechanism is a random process that depends only on the risky sex rate c_j and the relative proportion $p_j = n_j / \sum_{k=1}^m n_k$ of types j in the population (where again we have suppressed the time argument in $n_j(t)$ and $p_j(t)$, for these definitions must hold at all points in time). The mixing function is given by (Hethcote and Van Ark, 1987):

$$q_{ij} = q_j = \frac{c_j q_j}{E(c)}, \quad i, j = 1, 2, \dots, m, \quad (8)$$

where

$$E(c) = \sum_{j=1}^m c_j p_j \quad (9)$$

is the mean risky sex rate over the population. Equation (8) states that the likelihood of selecting a member of subpopulation j equals the fraction of all sexual partnerships in the population that involve type j men; note that

this probability is independent of i , as one would expect for random mixing. Equation (8) demonstrates a phenomenon known as random incidence. For a detailed discussion of random incidence within the context of AIDS modeling, see Kaplan (1989a).

4.2. Segregation

Again, assume that men are concerned with minimizing their perceived risk of infection. To do so, they attempt to select sex partners with the lowest possible risky sex rates. Segregation of the population by risky sex rate will occur when men have perfect information about prospective partners' risky sex rates, and when search costs are negligible. To see this, note that all men prefer partners with the lowest possible risky sex rate (say type 1). This is true of type 1's as well, however, so type 1's will match exclusively with other type 1's. Type 2's, upon realizing that type 1's are no longer available as sex partners (for type 1's match exclusively with type 1's), can do no better than to match with other type 2's. This segregation continues for all types of individuals. In general, for a type i individual, only type j individuals with the same or higher risky sex rates ($c_j \geq c_i$) are willing to be partners; and from this feasible set of partners, the type i man will prefer another type i , hence segregation by risky sex rate. Recall, however, that although men are assumed to know everyone's risky sex rate, we assume that men do not know how long prospective partners have been sexually active. The mixing function associated with segregation is merely the indicator function, that is

$$q_{ij} = 1_{(i=j)} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases} \quad i, j = 1, 2, \dots, m. \quad (10)$$

4.3. Bernoulli Noise (or Preferred Mixing)

The preferred mixing formulation to be described was proposed by Nold (1980) and has been employed by Hethcote and Van Ark (1987), Jacquez et al (1988), and Koopman et al (1988), though absent the following derivation. As an intermediate case between random mixing and segregation, imagine a world where all have imperfect information about prospective partners' risky sex rates. Here we consider the simplest model of imperfect information: every man has an apparent risky sex rate which everyone observes perfectly. With probability α , this apparent rate is guaranteed to equal the true rate. With probability $1-\alpha$, the apparent rate is a random incidence pick from the population of risky sex rates (i.e., c_j is observed with probability q_j as defined in equation (8); note that by chance, the random pick could result in the true rate). Hence, with probability α , the information is perfect, while with probability $1-\alpha$, the observed rate contains no information. Again, assume that search costs are negligible. In such a world, individuals will segregate based on observed rates following the reasoning of the segregation model. Both the random mixing model and the segregation model are special cases of Bernoulli noise with $\alpha=0$ and $\alpha=1$, respectively. With $0 < \alpha < 1$, while individuals still segregate, they do so on the basis of observed rates, allowing mixing among the true risky sex rates. Mixing according to Bernoulli noise can be thought of as an approximation to mixing that results from perfect knowledge of other, unmodeled, attributes such as attractiveness and familiarity.

To determine the mixing probabilities resulting from Bernoulli noise, define s_{ik} as the probability that a man with true risky sex rate c_i signals apparent risky sex rate c_k . In accordance with the assumptions postulated above, it is clear that

$$s_{ik} = \mathbf{a}1_{(i=k)} + (1 - \mathbf{a})q_k, \quad i, k = 1, 2, \dots, m. \quad (11)$$

Now define r_{kj} as the probability that a partner selected with an apparent risky sex rate equal to c_k has a true risky sex rate of c_j . Assuming random incidence within apparent risky sex classes (i.e., sex partners with equal apparent risky sex rates select each other at random), this probability is given by

$$r_{kj} = \frac{c_j P_j S_{jk}}{\sum_{\ell=1}^m c_\ell P_\ell S_{\ell k}}, \quad k, j = 1, 2, \dots, m. \quad (12)$$

The Bernoulli noise mixing probabilities q_{ij} are then given by

$$q_{ij} = \sum_{k=1}^m s_{ik} r_{kj} = \mathbf{a}^2 1_{(i=j)} + (1 - \mathbf{a}^2)q_j, \quad i, j = 1, 2, \dots, m, \quad (13)$$

with the last result following after tedious but straightforward algebra. Thus, the Bernoulli noise mixing probability function is simply a convex combination of the segregation and random mixing probability functions. Note again that $\alpha=1$ yields segregation while $\alpha=0$ yields random mixing.

4.4. Costly Search

Our final model of the partner selection process allows for mixing among men with different risky sex rates by relaxing the assumption that search costs are negligible. In contrast to the Bernoulli noise model (where information is imperfect and search costs are zero), here information is perfect but search costs are significant (but not large enough to qualify as enormous, which would force random mixing). In this case, an individual, if unsatisfied with a prospective partner due to the high risky sex rate, must consider the high cost of searching for a partner with a lower risky sex rate. In equilibrium, each individual will only be willing to search if the potential partner's risky sex rate is sufficiently higher than his own. For simplicity, we assume that $c_i=i$ for this model, and that individual preferences are such that a type i man will only accept partners with risky sex rates at most k above his own. Thus, a type i man will accept a partner of type j only if $j \leq i+k$. Similarly, type j 's will only accept type i 's if $i \leq j+k$. As a consequence, in equilibrium, type i 's will only be able to secure type j partners for $i-k \leq j \leq i+k$. Note that at $k=0$ the costly search model reduces to the segregation model. As k increases, the amount of mixing among individuals of different risky sex rates also increases.

To construct mixing probabilities consistent with these ideas, we first impose symmetry on the q_{ij} 's around the risky sex rate of the searching man, that is

$$q_{ij} = \begin{cases} q_{i,2i-j}, & \max(0, i-k) \leq j \leq \min(m, i+k) \\ 0 & \text{otherwise} \end{cases} \quad i, j = 1, 2, \dots, m. \quad (14)$$

This symmetry is somewhat similar to the use of Gaussian mixing functions by Hyman and Stanley (1988) We next impose $k(k+1)/2$ seed values for the mixing probabilities the values we specify are the k probabilities $q_{12}, q_{13}, \dots, q_{1,k+1}$, the $k-1$ probabilities $q_{24}, q_{25}, \dots, q_{2,k+2}, \dots$, and the probability $q_{k,2k}$. The imposition of these seed values, the symmetry of equation (14), the conservation law of equation (1), and the requirement that $\sum_{j=1}^m q_{ij} = 1$ for each i provide a set of linear equations which uniquely determine all of the mixing probabilities. One must be careful in choosing seed values however, for the solution of the equations just stated could yield values of q_{ij} which are negative; this is especially critical in dynamic simulations of the HIV epidemic, where the values of q_{ij} must be recalculated at each time instant

5. Epidemiological Implications of the Mixing Probabilities

In this section we will investigate the sensitivity of our model of HIV epidemics to the alternative mixing probabilities presented. To simplify matters, we will assume that the infectivities β_{ij} all equal the common value β ; this is a common assumption in the majority of published work to date (though one could certainly begin to assign differential values of β to different sexual practices such as unprotected or protected insertive versus receptive anal intercourse, for example). We will first present threshold conditions for the random mixing segregation, and Bernoulli noise models (the costly search model does not yield to more specific treatment than the general conditions Section 2 prescribes). Then we will study some numerical examples of random and nonrandom HIV epidemics generated from common assumptions regarding population immigration and emigration/death rates, AIDS incubation times, infectivity and the distribution of risky sex rates.

5.1. Threshold Condition for Random Mixing

The threshold condition for random mixing was presented by Anderson et al (1986) in terms of the reproductive rate of infection. In our notation, their result may be stated as follows: an HIV epidemic under random mixing occurs if the reproductive rate of infection R_0 exceeds unity, where

$$R_0 \equiv \frac{E(c^2)}{E(c)} bD \quad (15)$$

and $E(c^2) = \sum_{i=1}^m c_i^2 p_i$. To obtain this result from our approach in Section 3, note that the appropriate value of q_{ij}^* to use in equation (5) is

$$q_{ij}^* = q_j^* \equiv \frac{N_j c_j}{\sum_{k=1}^m N_k c_k}, \quad j = 1, 2, \dots, m. \quad (16)$$

Inserting (16) into (5), and assuming that $\beta_{ij}=\beta$ for every i, j pair, we obtain the equation

$$T_i = c_i bDK, \quad i = 1, 2, \dots, m. \quad (17)$$

where

$$K = \sum_{j=1}^m q_j^* (1 + T_j) = \sum_{j=1}^m q_j^* (1 + c_j bDK) = \frac{1}{(1 - R_0)}, \quad (18)$$

the last equality following by direct substitution for q_j^* . The theory of Section 2 states that an epidemic occurs if and only if $T_i < 0$ for some i ; from equations (17) and (18), it is clear that $T_i < 0$ for every i if and only if $R_0 > 1$.

5.2. Thresholds for Segregation

Having derived the epidemic threshold for random mixing, the analysis for segregation is automatic; for any particular segregated subpopulation i can be viewed as a randomly mixing population with mean risky sex rate c_i and zero variance. Thus, an epidemic will occur in subpopulation i if and only if

$$R_{0i} \equiv c_i bD > 1, \quad i = 1, 2, \dots, m, \quad (19)$$

and an epidemic will occur somewhere in the population if and only if $R_{0m} > 1$ (assuming c_m is the largest risky sex rate). Already we have the possibility that a particular distribution of risky sex rates might not sustain an HIV epidemic under random mixing (for R_0 of equation (15) might fall below unity), yet this same distribution of risky sex rates could sustain an epidemic under segregation (for R_{0m} could exceed unity even though R_0 does not).

5.3. Threshold for Bernoulli Noise

Recall from equation (16) the definition of the Bernoulli noise mixing probabilities. Given this simple form, one might expect that the threshold condition for a Bernoulli noise epidemic is a simple combination of the thresholds for random mixing and segregation. This is not the case, as we will now show.

First, consider the largest risky sex rate c_m . If $\alpha^2 R_{0m}$ exceeds unity, an epidemic will develop in subgroup m , for men in subgroup m mix among themselves sufficiently to generate a segregated epidemic even after discounting sexual contacts between men in subgroup m and other subgroups. In this instance, an epidemic will develop among all subgroups, as all subgroups communicate for $0 < \alpha < 1$. Thus, a sufficient condition for a Bernoulli noise epidemic is that $\alpha^2 R_{0m}$ exceeds unity.

Now suppose that $\alpha^2 R_{0m}$ is less than one. To use equation (5), note that the appropriate value of q_{ij}^* ; equals

$$q_{ij}^* = a^2 1_{(i=j)} + (1 - a^2) q_j^*, \quad i, j = 1, 2, \dots, m, \quad (20)$$

where q_j^* ; is given in equation (16). Use of equation (5) then implies that the threshold condition (when $\alpha^2 R_{0m} < 1$) for a Bernoulli noise epidemic is crossed when

$$(1 - \alpha^2) \sum_{k=1}^m q_k^* R_{0k} / (1 - \alpha^2 R_{0k}) > 1. \quad (21)$$

Note that when $\alpha=0$, (21) reduces to $R_0 > 1$ as expected for random mixing. Also, as the left hand side of (21) exceeds $(1 - \alpha^2) R_0$, another sufficient condition for a Bernoulli noise epidemic to occur is $(1 - \alpha^2) R_0 > 1$ (this mirrors the sufficient condition $\alpha^2 R_{0m} > 1$ discussed earlier).

The threshold in (21) has an interesting implication. It is entirely possible that a given distribution of risky sex rate could yield $R_0 < 1$ under random mixing, but for some value of α , the threshold in (21) could be satisfied. As random mixing corresponds to no information about the risky sex rates of prospective partners, while Bernoulli noise implies some information (as measured by α), imperfect information can be worse than no information. Worded differently, a little information can be a dangerous thing. This result differs from our earlier comparison between random mixing and segregation, for if the threshold in (21) is crossed, then all subgroups will experience an HIV epidemic (as opposed to only some of the subgroups as is the case with segregation). Of course, this result is a theoretical possibility; for risky sex distributions which roughly describe the sexual activity of gay men, the severity of HIV epidemics under random mixing is typically harsher than that experienced under any of our nonrandom mixing models, as we will now illustrate.

5.4. Numerical Examples of Nonrandom HIV Epidemics

To simulate HIV epidemics under the various mixing schemes discussed, we make the simplifying assumption that the AIDS incubation time is exponentially distributed with mean $1/\gamma$. Though the exponential is not descriptively accurate, its use simplifies our computations without distorting the impact of the various mixing patterns we consider. Define $X_i(t)$ and $Y_i(t)$ as the number of uninfected and infected men in subpopulation i at time t , $X_i(t) + Y_i(t) = n_i(t)$. Further, define $\pi_i(t) = Y_i(t)/n_i(t)$ as the prevalence of infection in the i^{th} subpopulation. Our assumptions imply the following differential equations:

$$\frac{dX_i(t)}{dt} = N_i - mX_i(t) - (1 - p_i(t)) \sum_{j=1}^m c_j Y_j(t) q_{ji}(t) b_{ji}, \quad i = 1, 2, \dots, m, \quad (22)$$

and

$$\frac{dY_i(t)}{dt} = (1 - p_i(t)) \sum_{j=1}^m c_j Y_j(t) q_{ji}(t) b_{ji} - (m + g) Y_i(t), \quad i = 1, 2, \dots, m. \quad (23)$$

In the absence of HIV/AIDS, we assume that sex lives average 32 years, thus $\mu = 1/32$ per man per year (as assumed by Anderson et al, 1986 and Kaplan, 1989a). The incubation time for AIDS is assumed to average 8 years (see Medley et al, 1987 and Lui, Darrow and Rutherford, 1988) so $\gamma = 1/8$ per man per year. The infectivity β is set equal to .075 per partnership (this falls within the range of infectivities for receptive anal intercourse

estimated by Grant, Wiley and Winkelstein, 1987 and Kaplan, 1989a). Finally, the immigration rates N_i for men with risky sex rates of $c_i=i$ partners per year are assumed to follow

$$N_i = 3125 \frac{11-i}{55}, \quad i = 1, 2, \dots, 10, \quad (24)$$

In effect we are approximating heterogeneity in risky sex rates by dividing the population into 10 different subpopulations. In the absence of HIV, the sexually active population would reach an equilibrium of 100,000 men. The mean risky sex rate in the absence of HIV equals 4 partners per year, while the ratio $E(c^2)/E(c)$ equals 5.5 partners per year. Though less variable than the risky sex distributions considered by Anderson et al (1986) and Kaplan (1989a), the mean risky sex rate of 4 is comparable to the average of 5 partners per year employed by these researchers.

To introduce HIV to the population, we establish the following initial conditions:

$$Y_i(0) = n_i(0)q^{i-10} / n_{10}(0) \quad i = 1, 2, \dots, 10, \quad (25)$$

where $n_i(0)=N_i/\mu$. Note that $Y_{10}(0)=1$; equation (25) introduces a single infected man into the most active subpopulation and essentially introduces a trace of infection into each of the other subpopulations. Throughout our numerical examples, we set $\theta=5$.

Figure 1 plots the total number of infected men in the population (i.e., $\sum_{i=1}^{10} Y_i(t)$) versus time for Bernoulli noise epidemics where the strength of the information signal (A. measured by α^2) is varied from 0 (i.e., random mixing) to 1 (i.e., segregation). The reproductive rate of infection for the random mixing case, R_0 (equation (15)), equals 2.64, thus the ensuing epidemic is hardly a surprise. For the parameters of this example, the Bernoulli noise epidemics with $\alpha^2=.25, .5$ and $.75$ closely mimic the random mixing epidemic, with the time until peak infection dropping slightly as α increases. However, the segregation epidemic resulting when $\alpha=1$ has a distinctly different pattern. The epidemic moves in waves as each of the subpopulations from 10 through 3 reach their peak levels of infection at different times (after the 100 years simulated here, subpopulations 3 and 4 had yet to reach their peak levels of infection). This is due to two factors: the differences in risky sex rates for the different subpopulations, and the different initial conditions (as stated in equation (25)). The initial conditions in this instance mimic the phased introduction of HIV into the various subpopulations. What is clear from Figure 1, however, is that the random mixing epidemic is the most severe of those considered. These limited simulations suggest that the random mixing model may provide a conservative tool for policy analysts.

It is possible, however, for a nonrandom mixing epidemic to occur when $R_0 \leq 1$. For example, suppose that the infectivity was reduced from .075 to .025 due to some control measure (perhaps due to the sporadic use of condoms). In this case, no random mixing epidemic would ensue, for the reproductive rate R_0 of equation (15) would equal .88, which is less than unity. However, Bernoulli noise epidemics are possible: the threshold condition (21) implies that HIV infection will endure if $\alpha^2 > .4427$. The time necessary for such epidemics to develop, however, is long indeed. For example, after a century of HIV infection following the initial conditions of equation (25), about 100 of 99,800 men would be infected with $\alpha^2=.75$; segregation yields 585 infections out of 98,600 men after the same time period. Utilization of the random mixing model to determine that the control measure would be effective (as $R_0 < 1$) does not lead to disaster in the event that the population in question really mixes in accordance with Bernoulli noise. The resulting prevalence of infection in this latter case remains low:

no higher than $585/98,600 \approx .006$ after 100 years. Such low levels of infection over such a long time period cannot be considered serious. Though no hope exists for a rapid medical solution to HIV infection, the same cannot be said for a time frame of a century! Also, behavioral change will surely obviate these results with the passage of time.

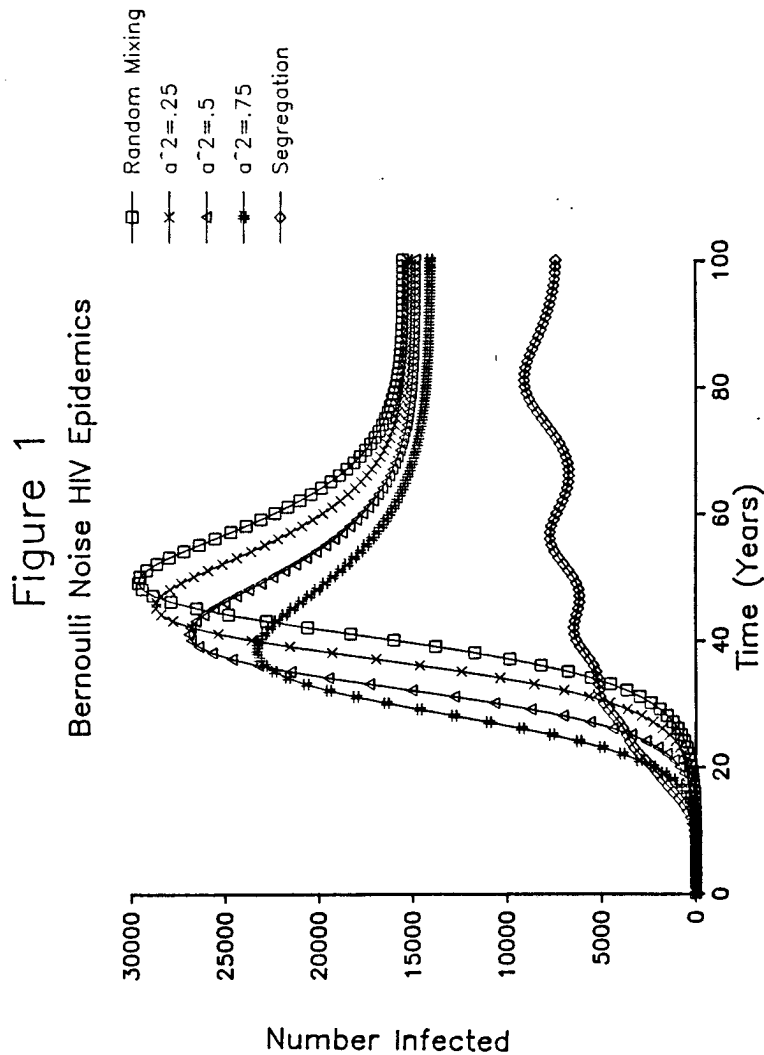
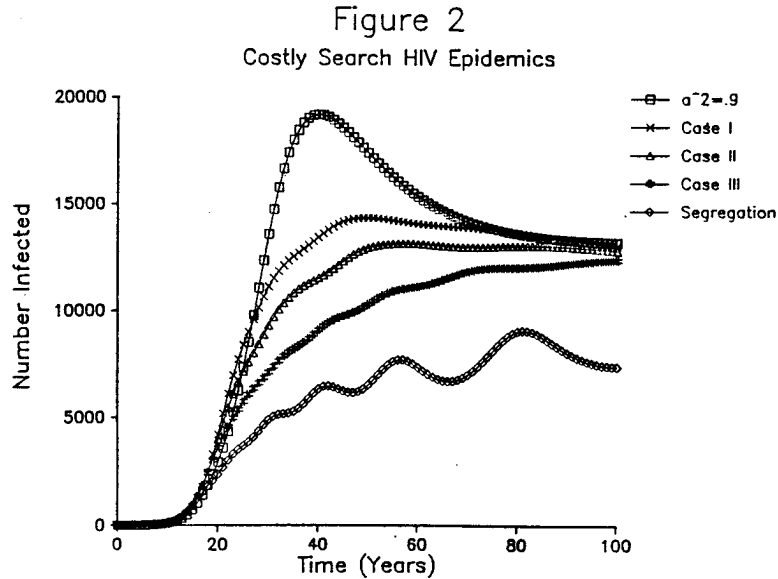


Figure 2 presents three costly search epidemics, along with the segregation epidemic and a Bernoulli noise epidemic with $\alpha^2 = .9$ for comparison. In Case I, $q_{12}(t) = .1$ and $q_{13}(t) = q_{24}(t) = .05$; $q_{12}(t) = .075$ and $q_{13}(t) = q_{24}(t) = .025$. In both of these cases, a man is restricted to choosing partners with risky sex rates that differ by, at most, two from his own. Case III is even more stringent; a man is restricted to partners with risky sex rates that differ by, at most, one from his own. In this case, $q_{12}(t) = .1$. Generally speaking, these mixing functions possess modes with probability mass between .5 and .9 at $q_{ij}(t)$ for all time periods and most subpopulations (though $q_{10,10}(t)$ is not

modal for subgroup 10 for some time periods in these simulations). Note that the epidemics resulting from the costly search mixing functions have been flattened relatively to Bernoulli noise; there is no rapid rise to peak infection followed by a quick retreat towards steady state. Indeed, the costly search epidemics appear to be smoothed versions of the segregation epidemic, though higher levels of infection are reached. However, the long run levels of HIV infection for the costly search models appear comparable to the Bernoulli noise and random mixing models. Again, random mixing appears to be a worst case.



6. Concluding Remarks

This paper has formulated a macromodel of nonrandom mixing, two different approaches for deriving threshold conditions for this model, and discussed new approaches to deriving mixing probabilities. Four mixing scenarios were discussed: random mixing segregation, Bernoulli noise, and costly search. Although it is possible for nonrandom models to produce more serious epidemics than random mixing dictates, we found that epidemiological environments roughly descriptive of the AIDS epidemic produce worse epidemics for random mixing than for the nonrandom models studied.

There is little doubt that, from a purely descriptive point of view, random mixing does not describe human behavior; by implication, some variant of nonrandom mixing does. The important question, however, is whether or not random mixing models of HIV infection can seriously mislead policy analysts working within a fixed epidemiological environment (that is, given a distribution of risky sex rates, immigration and emigration/death rates, infectivities and an incubation time distribution).

Our limited experience suggests that for a given epidemiological environment, either random mixing serves as a worst case over nonrandom mixing scenarios (though the bias does not appear to be too extreme), or when random mixing epidemics are not severe, neither are nonrandom epidemics. Thus, having derived a variety of nonrandom mixing models, an important question remains: for a fixed epidemiological environment, can one

have a significant nonrandom mixing epidemic (i.e., high prevalence of infection over a short time scale, say 20-30 years) when random mixing does not produce a significant epidemic for the same epidemiological parameters? To our knowledge, this question remains open.

Acknowledgement

We wish to thank Carlo Castillo-Chavez and the six anonymous referees their suggestions have strengthened this paper.

References

- Abramson, P.R. and B. Rothschild (1988). Sex, drugs and matrices: mathematical prediction of HIV infection. J. Sex Res., 25, 106-122.
- Anderson, R.M. (1988). The Epidemiology of HIV Infection: Variable incubation plus infectious periods and heterogeneity in sexual activity. J. Roy. Stat. Soc. (Ser. A), 151, 66-93
- Anderson, R.M., G. F. Medley, R.M. May, and A.M. Johnson (1986). A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. IMA J. Math. Appl. Med. Bio., 3, 229-263.
- Blythe, S.P. and C. Castillo-Chavez (1989). Like-with-like preference and sexual mixing models. Math. Biosci. (in press).
- Brookmeyer, R. and M., Gail (1988). A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. JASA, 83, 301-308.
- CDC (1989). HIV/AIDS Surveillance Report. Centers for Disease Control, Public Health Service, U.S. Department of Health and Human Services, May 1989, 1-16.
- Curran, J.W., H.W. Jaffe, A.M. Hardy, W. M. Morgan, R.M. Selki, and T.J. Dondero (1988). Epidemiology of HIV infection and AIDS in the United States. Science, 239, 610-616.
- Grant, R.M., J.A. Wiley, and W. Winkelstein (1987). Infectivity of the human immunodeficiency virus: estimates from a prospective study of homosexual men. J. Inf. Dis., 156, 189-193.
- Hethcote, H.W. and J.W. Van Ark (1987). Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs. Math. Biosci., 84, 85-118.
- Hethcote, H. W. and J. A. Yorke (1984). Gonorrhea, transmission dynamics, and control. Lecture Notes in Biomathematics 56, Springer-Verlag, Berlin.
- Hyman, J.M. and E.A. Stanley (1988). Using mathematical models to understand the AIDS epidemic. Math. Biosci., 90, 415 - 473.
- Isham, V. (1988). Mathematical modeling of the transmission dynamics of HIV infections and AIDS: a review. J. Roy. Stat. Soc. (Ser. A), 151, 5-30.
- Jacquez, J.A., C.P. Simon, J. Koopman, L. Sattenspiel, and T. Perry, (1988). Modeling and analyzing HIV transmission: the effect of contact patterns. Math. Biosci., 92, 119-199.
- Kaplan, E. H. (1989a). What are the risks of risky sex? Modeling the AIDS epidemic. Op. Res., 37, 198-209.
- Kaplan, E.H. (1989b). Needles that kill: modeling human immunodeficiency virus transmission via shared drug injection equipment in shooting galleries. Rev. Inf. Dis., 11, 289-298.
- Kaplan, E.H. and P.R. Abramson (1989). So what if the program ain't perfect? A mathematical model of AIDS education. Eval. Rev., 13, 107-122.
- Koopman, J., C. Simon, J. Jacquez, J. Joseph, L. Sattenspiel, and T. Park, (1988). Sexual partner selectiveness effects of homosexual HIV transmission dynamics. JAIDS, 1, 486-504.
- Lui, K-J, W.W. Darrow, and G.W. Rutherford, III (1988). A model-based estimate of the mean incubation period for AIDS in homosexual men. Science, 240, 1333-1335.
- May, R.M. and R.M. Anderson (1987). Transmission dynamics of HIV infection. Nature, 326, 137-142.
- May, R.M. and R.M. Anderson (1988). Transmission dynamics of human immunodeficiency virus (HIV). Phil. Trans. Roy. Soc. (in press).
- Medley, G.F., R.M. Anderson, D.R. Cox, and L. Billard (1987). Incubation period of AIDS in patients infected via blood transfusion. Nature, 328, 719-721.
- Nold, A. (1980). Heterogeneity in diseases-transmission modeling. Math. Biosci. 52, 227-240.
- Strang, G. (1980). Linear Algebra and its Applications. Academic Press, New York.
- Wiley, J.A., S. Herschkorn, and N. Padian (1989). Heterogeneity in the probability of HIV transmission per sexual contact. Stat. Med., 8, 93-102.