

LEGIBILITY NOTICE

A major purpose of the Technical Information Center is to provide the broadest dissemination possible of information contained in DOE's Research and Development Reports to business, industry, the academic community, and federal, state and local governments.

Although a small portion of this report is not reproducible, it is being made available to expedite the availability of information on the research discussed herein.

TITLE: A MATHEMATICAL MODEL OF THE SPREAD OF THE AIDS VIRUS

LA-UR--87-1572

DE87 010112

AUTHOR(S): J. M. Hyman and E. A. Stanley

SUBMITTED TO: Third International Conference on AIDS,
Washington, DC, June 1-5, 1987

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes.

The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

A MATHEMATICAL MODEL OF THE SPREAD OF THE AIDS VIRUS

Prepared by
J. M. Hyman and E. A. Stanley

*Center for Nonlinear Studies
Theoretical Division, B284
Los Alamos National Laboratory
Los Alamos, NM 87545*

ABSTRACT

A mathematical computer model of the spread of the AIDS epidemic in the United States is being developed at Los Alamos National Laboratory. This model predicts the spreading of the HIV infection, and subsequent development of clinical AIDS in various population groups. These groups are chosen according to age, frequency and type of sexual contact, population density, and region of the country. Type of sexual contact includes not only the heterosexual, homosexual differentiation but also repeated contacts with such primary partners as spouses. In conjunction with the computer model, we are developing a database containing relevant information on the natural history of the viral infection, the prevalence of the infection and of clinical AIDS in the population, the distribution of people into sexual behavior groups as a function of age and information on interregional contacts. The effects of variable infectiousness and sexual activity during the long period from infection to disease are found to have a major impact on the predictions of the model.

Research Investigators include: *W. A. Beyer, S. R. Booth, S. A. Colgate, R. H. Drake, P. S. Hagan, J. M. Hyman, S. P. Layne, T. G. Marr, J.C Peterson, C. B. Qualls, R. T. Pfaff, E. A. Stanley and L. A. Waller*

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED



OVERVIEW OF THE PROPOSAL

A. WHAT WE INTEND TO DO

This proposal is to develop and investigate a computer model to predict the future course of the AIDS epidemic in the United States. This model will describe spreading of the HIV infection in the different age groups and regions of the country, and from groups engaging in extremely high risk behaviors to those engaging in lower risk behaviors, and estimate the subsequent AIDS caseload. The future spread of the virus will most likely be through sexual contact and drug needle sharing, these are accounted for in great detail in the model. (The transmission in Africa would require also including blood transfusions.) These predictions could be used to analyze the medical care costs and national productivity losses due to illness and death.

Current predictions of the spread of the AIDS epidemic are based on simple exponential or polynomial extrapolations. It is inevitable that behavior changes and medical advances will greatly affect the future course of the epidemic. The changes will be highly nonlinear functions of the parameter values and at times may even lead to changes that are counter to intuition and to simple extrapolation predictions. The predictions must realistically take these changes into account. The usual curve fitting methods are not reliable for complicated nonlinear behavior, especially if long time scales such as the virus incubation period are crucial. The mathematical model predictions of these counter-intuitive mechanisms may greatly improve our understanding of the observations.

A central goal of the model is to identify which variables the predictions are most sensitive to. These are the most important variables to collect detailed information on in future epidemiological studies. The model will also allow health officials to evaluate the effect of different strategies (education, use of condoms, spermicides, change in sexual behavior, ...) on the spread of HIV. Using mathematical models to evaluate strategies has been applied successfully in other epidemics [Martini, Heathcote and York]; most dramatically in predicting the influence of mosquito population on the spread of malaria [Ross].

To achieve this goal we plan to develop as detailed and scientifically sound an epidemiology model for the spread of the HIV infection as possible. The current data and computer power preclude actually using this full model. However, the full model can be used to mathematically and computationally determine which are the

most appropriate variables and simplified models to predict the spread of the epidemic.

Our model is not really a single model, but rather a sequence of models, of increasing complexity. For example, one version may neglect age dependence, while another may incorporate it. All of these models fit into the same general framework, and we are developing a computer program which allows the models to be easily changed to incorporate new ideas or to compare one version with another.

We will use the model predictions to increase the awareness and make recommendations to policy makers by identifying the most effective measures that could be taken to curb the spread of the HIV infection.

B. WHY THIS IS IMPORTANT

Our primary objective is to quickly gain the capability to accurately predict the future incidence of clinical AIDS. Unless significant medical breakthroughs change the course of the epidemic, the large number of currently infected people (1.5 - 2 million in the US) will increase and most will eventually convert to AIDS. As the epidemic progresses, the need for accurate predictions of health care needs and of the economic impact will increase.

It is likely that large numbers of Americans will become ill and die from this disease before either a vaccine or an effective treatment becomes available [Hahn, et. al.]. Most of these people will be in the 20-40 age group [Burke, Francis and Chin]. Both the loss of productivity and the burden on the health care system will be enormous. In the long run, the military preparedness of the country may be affected. The more accurately we can estimate the extent and distribution of the epidemic, the better prepared the nation can be to deal with the care of those affected. We will use our predictions to work in conjunction with other efforts to analyze the health care costs and the economic losses due to the premature deaths of so many young people.

The spreading of HIV by heterosexual contacts can already be seen in the most recent figures from CDC. In the adult black population, heterosexually transmitted AIDS already accounts for 12% of all cases. In the tests of military recruits the fractions of infected males and of females are nearly equal [Burke].

In our preliminary calculations, the safety of the heterosexual community is largely an illusion based on the different time scales associated with different sexual behavior patterns. In these runs the heterosexual (less active) susceptibles mimic

the current homosexual trends with a time lag of only a few years. This is true, even though only a fraction of 1% of the heterosexual population is currently infected.

Besides the obvious desire to accurately predict future numbers of infected, ill, and dead we are developing a tool which will allow investigation of the effects of behavioral changes, treatment methods, and vaccines on the future course of the epidemic. This will allow investigators to answer many questions. For example, one can assume increased condom use by people in a targeted age group and region and determine how much that increased use will slow the local course of the epidemic. This would then help authorities to decide if it is more effective to encourage condom use in that group, than to use another strategy. As another example, eventually a partially effective vaccine will be developed. Somehow it must be ascertained which persons should be vaccinated. The model can be used to understand the effect of vaccinating each group on the spreading of the epidemic.

C. WHY US?

Los Alamos is particularly well-equipped to develop the HIV spreading model. We have already assembled a broad-based, high-caliber scientific team which includes scientists with previous experience in modeling epidemics, numerical analysis, computer science and economic analysis. These team members are active research scientists and, although most are not currently supported for their epidemiology efforts, they are hard at work primarily due to their concern about the future course of the epidemic.

In a scientifically sound epidemiological model, the mathematical aspects are crucial to predicting the future course of the infection. The variables considered in the model are selected according to their influence on the predictions. Defining an accurate model is an exercise in applying a well based scientific methodology to understand the underlying transmission mechanisms.

These predictions may not always be intuitive. For example, the past underpredictions of how many infected individuals will develop AIDS resulted from a counter intuitive mechanism and delayed public awareness of the extent of the problem. If most of the AIDS cases diagnosed today were infected 3-4 years ago (Fig. 1a) what does this say about the average time to convert to AIDS? If the HIV infected population was constant the answer would be 3-4 years, but it's not. The HIV infected population is changing rapidly. In Fig. 1b we give an example where the HIV infected population is doubling every year and the average time to convert to AIDS is 7-8 years. The distribution of patients currently developing AIDS in Fig. 1a

is the product of the two curves in Fig. 1b, and hence, is highly skewed toward the early conversion times. In this simplified example, it is clear how the model can increase our understanding of the observations and improve our predictions.

Even our current, relatively crude, computer models require hundreds of thousands of differential equations to be solved. As the computer model becomes more complex, supercomputers will be necessary for accurate detailed predictions. The Los Alamos Computer Facility contains the most powerful collection of scientific supercomputers in the world. These computers will allow us to investigate an extremely detailed model, and to maintain and analyze a large data base which will give us the information necessary to make this detailed model work.

Although the economic and social impact of the epidemic is not a primary goal of this proposal, it will be of utmost importance. The best way to ensure that our transmission model predictions are known and understood by the people making cost estimates is to have local scientists with experience in cost benefit analysis involved. Los Alamos has some expertise in this field [references], and we can use this expertise to ensure that our results are used appropriately.

As the predictions are obtained, the Los Alamos statisticians and economists can work with other agencies to use the predictions to analyze and estimate the expected economic costs. These costs result both from medical treatment of AIDS patients and from productivity losses due to their illness and premature death. Medical system costs include not only direct treatment, but also research expenditures and problems created in the insurance and health care systems. The productivity losses and economic impact due to losses in expected earnings may prove to be immense. Most of those affected are young, with many productive years ahead of them. Our economic analysts will be able to do some preliminary estimates of the medical care costs and productivity losses, but they expect to primarily serve as a conduit to pass our results on to these other groups.

D. OUR BASIC TIME SCALE

Our model for the spread of the HIV infection is progressing rapidly. We expect to have a preliminary version of the model running by early June. The model will use the data we are collecting from the CDC, NIH, the armed services, state and local health officials, and other groups studying the AIDS epidemic, on the progression of the infection in individuals, the transmissibility of the disease and the prevalence of the virus in different segments of the population. This information will be organized into a data base and used to evaluate parameters of the model.

We are also gathering sociological information on age specific sexual activity and drug use which will allow us to gradually improve the basic model. The model will be carefully monitored to ensure consistency with known information on the progression of the epidemic and to gain a thorough understanding of the effects of the new changes. Within the first year, the model will mimic the major transmission lines of the disease and be used to predict the future spreading of the HIV infection.

The current poor data base, on which the model depends, leads to large error bounds on the predictions. As the quality and quantity of the data improves, so should the predictions. However, just because the optimal data doesn't exist the model development should not be delayed. The final model will take several man-years to implement and can be used to improve the internal consistency of the data. The model can also identify anomalies, suspected "bad" data points and what data is most needed to improve the predictions. We are maintaining contact with the other AIDS modeling efforts. Because of the uncertainties in the data and the highly nonlinear aspects of the problem we expect the different forecasting techniques to give qualitatively similar predictions but significant differences in the details.

As this model is developed, we will interact continuously with the various groups which are collecting and analyzing data on the modes of transmission, the prevalence in the population, and the etiology of the HIV viral infection. The model development will raise new questions at each stage and new information will become available as more and more studies are undertaken. We expect that our questions will help to focus some of these studies on aspects of the epidemic which must be understood in order to make accurate predictions. By requiring the data to be self-consistent, the model can estimate unknown data and parameters from known ones. For example, how many people were infected with the HIV as a function of the known AIDS cases. Once we feel that the model is sufficiently well developed, we intend to make the computer program available to other scientists to conduct their own investigations with it.

We are currently working on the model development on four fronts. First we are talking to clinical epidemiologists and searching through the clinical literature in order to gain the best idea possible of the etiology of the HIV infection: the infectiousness of different routes; the variation of infectivity as the infection progresses; the distribution of times after infection at which AIDS occurs and at which death occurs; how these variables depend on the general health of an

individual, etc. Second, we are investigating the effects of adding various factors one or two at a time, such as variable infectivity, to the Anderson, et al. homosexual model. Third, we are writing the general computer model described above. Finally, we are talking to demographers and sociologists, as well as AIDS investigators, in order to accurately model sexual behavior and drug use.

SUMMARY

Major medical advances are required before an effective antiviral therapy is developed or an effective vaccine will be widely available. Thus, we have to prepare for a long battle against the spread of the HIV infection. The model will help us gain insight into how the epidemic is developing and allow us to visualize the future. The formal mathematical model will allow computer simulations of the large scale dynamics of the epidemic. The complex interactions of the various competing forces can be singled out and individually studied to improve our understanding of the essential relationships between the social and biological mechanisms that influence the spread of the disease. Many of the computer experiments, run to gain insight into the behavior of the epidemic, would not be possible or ethical to perform in human populations. The influence and sensitivity of various factors on the spread of the epidemic can be ascertained. Those factors that are most crucial will be identified and used to set priorities in the research and target the most effective means to slow the spread. The model will be a tool where we can run computer experiments to predict the outcome of different approaches. These experiments can save time, resources and lives. It will allow us to predict what the future really will be and act as a control group for true experimental situations.

TASKS

1. Develop a realistic comprehensive model for heterosexual spreading of the HIV infection without undue concern of the currently available data or computational complexity of the model. This comprehensive model will be used to determine the appropriate computationally solvable submodels on which the predictions will be based. The model will describe transmission by both sexual contact and sharing of needles. Sexual transmission is from male to female, female to male and male to male. Contacts occur between people living in different regions and between different age groups. Numbers of children infected perinatally as well as the numbers of clinically ill will be determined from the

numbers infected for a given length of time in the appropriate age groups. Development of the model involves a number of processes:

- a. defining the general integral—partial differential equation framework
 - b. understanding and modelling the contact rates between various groups.
 - c. simplifying the model to forms which the computer can solve in reasonable time periods, using estimates of parameter values. (E. A. Stanley, T. G. Marr, J. M. Hyman, P. S. Hagan).
2. Implement a computer code to solve the model. This code will be based on the general framework of the model, so that it is sufficiently flexible to be changed as the model changes. It must be easy to use. The large number of variables involved make it difficult to visualize both the input and the results. Thus, a flexible input and output interface, coupled to good graphics software, will be developed. (J. M. Hyman, R. T. Pfaff, J. Peterson)
 3. Obtain data on the AIDS epidemic. This includes data on seroprevalence, numbers of clinical AIDS cases, and from clinical cohort studies. Clinical cohort study data will give information on the distribution of times between infection and development of clinical AIDS and between AIDS and death, on the infectivity of individuals as infection progresses, and on sexual behavior patterns, all of which depend on age and general health. This data will be continuously updated. Obtaining this information involves interacting with agencies across the country, including the various agencies of the NIH, the armed services, hospitals, and public health services. (E. A. Stanley, S. P. Layne, S. A. Colgate)
 4. Organize and do statistical analysis of the data so that it is in a form which the computer model can use. Understand the biases in each sample and attempt to account for them. (C. B. Qualls, T. G. Marr, W. A. Beyer, L. A. Waller)
 5. Understand sexual and needle-sharing behavior as it varies with age, and as it is affected by knowledge of infection and fear of contracting the virus. This involves interacting with sociologists and clinical workers. The numbers of partners and frequency and type of sexual contact with the same partner are all important to understanding the transmission lines of this disease. For the age structured model, frequencies of contact between age groups need to be estimated. (T. G. Marr, E. A. Stanley, L. A. Waller)
 6. Explore the behavior of the model. The sensitivity of parameters and any inconsistencies with the data must be understood to ensure that we have captured the essential transmission mechanisms. If the model can be simplified without loss of important information, this should be done. Explaining our results to clinical

investigators and obtaining feedback from them will be an important aspect of this stage. Inconsistencies in the data and the sensitivity of the model to parameter changes will suggest case studies that should be done. These suggestions will be passed on to other agencies. The model will be continuously revised to keep up with current knowledge. (S. A. Colgate, P. S. Hagan, J. M. Hyman, T. G. Marr, E. A. Stanley, L. A. Waller)

7. Make predictions with the model. Compare predictions based on current behavior with those based on behavior changes wrought by education and fear of contagion. (S. A. Colgate, P. S. Hagan, J. M. Hyman, T. G. Marr, E. A. Stanley, L. A. Waller)

8. Develop models to estimate health care costs and productivity losses due to the illnesses and death caused by AIDS. Use the above predictions to make these estimates. Interact with NIH, the CDC, armed forces and others to understand the future impact of this epidemic on the health care system and on the framework of society. (S. R. Booth, R. H. Drake)

9. Release the model to other scientists for their own investigation of its predictions and of the effects of treatment programs and vaccines. This release will occur only after the model is thoroughly understood and is known to be robust.

ADDENDUM I

DESCRIPTION OF THE TRANSMISSION MODEL

INTRODUCTION

HIV is primarily transmitted through sexual contact (M-F, M-M), sharing of hypodermic needles and exposure to infected blood either perinatally or through blood transfusions. It is not transmitted by nonsexual daily contacts, even though the virus has been isolated from almost every body fluid [Fischl, et. al.]. The infection risk to an individual depends both on the behavior of the individual and on the prevalence of infection in the groups with which the individual has sexual contacts or shares needles. This prevalence varies between regions and age groups, as well as between behavioral risk groups. Behavioral risk groups are defined not only by the frequency of changing homosexual and heterosexual partners and of sharing needles, but also by the manner in which the partners are chosen. An individual is more likely to be infected if he or she

- ▶ has multiple sexual partners
- ▶ has sexual partners in a high risk group

- ▶ lives in a highly populated area
- ▶ lives in the New York City, Washington, DC, San Francisco or Los Angeles areas
- ▶ shares needles when using drugs
- ▶ is between 25-35 years of age

We are developing a fairly general framework for our model of the transmission of the HIV infection. This framework attempts to include all of the risk factors which we foresee as being important to the epidemic, and some which will eventually be found to be unimportant. Posing the model in this general manner has several advantages. It allows us to develop a computer program which is sufficiently flexible to be able to switch from the most simplified to the most complex of models. It allows us to ask questions about the behaviors of different groups of people and to attempt to ascertain which behaviors are essential to understanding the spreading of the epidemic.

The portion of the male and female population that engages in behaviors which put them at risk for HIV, namely, nonmonogamous sexual contact and needle-sharing drug use, is divided up according to their risk behaviors and the manner by which they choose partners. Susceptible persons are infected through contacts with infected persons, and infected persons develop clinical AIDS either Kaposi's sarcoma or opportunistic infections, at a rate which depends on the length of time since infection began. AIDS patients subsequently die at a rate which depends on the length of time since AIDS developed, and on the type of infection (either KS or opportunistic).

It is assumed that once a person becomes infected, he remains infected for life. This one-way migration of susceptibles to infected is due to the chromosomal integration of the proviral DNA into the host cell. Numbers of children infected perinatally are estimated indirectly from the numbers of infected adults.

In its most detailed form, the model, which is described in detail in subsequent sections, contains too many variables to be solved numerically on even the largest and most advanced computers. Even if it were possible to solve the system, not enough will ever be known about human behavior to supply the necessary information to the program, and any results would be so complex as to preclude understanding. Instead, simplified submodels must be distilled out of the full model; and their behavior investigated. This will allow understanding of the interactions of different factors in the spreading of the AIDS virus. For example, in order to comprehend how well the infectivity profile (infectiousness as time since

infection) must be measured, one can look at the sensitivity of a very simple model to variations in the profile. Such a model can lump age groups, regions and even ignore sex, but cannot ignore all nonheterogeneities in sexual partner choices. On the other hand, if we wish to understand how difference in age may delay spreading into one age group from another, then we cannot ignore age-structured behavior.

By requiring the data to be self-consistent, the model can estimate unknown data and parameters from known ones. For example, how many people were infected with the HIV as a function of the known AIDS cases. As we explore various sub-models, we will build up a picture of interactions which will allow us to make estimates in the global model leading to major simplifications. We will also raise questions and insights which will spark investigators into reformulating their research, and add to the general understanding of this epidemic.

Thus our model is not really a single model, but rather a collection of models, with varying complexity, each emphasizing different medical and social factors. For example, one version may neglect age dependence, while another may incorporate it. All of these models fit into the same general framework, and we are developing a computer program which deals with the framework and thus allows the model to be easily changed to incorporate new ideas or to compare one version with another. Also, the optimal approximating submodel of the full model during the growth of the infection may not be the optimal approximation when the epidemic starts to stabilize. The best approximation depends upon the state of the system and the better computer models dynamically adapt their resolution of the full model as the solution evolves. As more detailed statistical information becomes available, the more sophisticated models will become increasingly more appropriate and accurate.

RISK FACTORS

In contrast to our current understanding of the transmission of malaria [Ross], measles [Dietz and Schenzle], rubella [?], rabies [Murray, et al.] and many other diseases [Anderson and May 1982], little is known about modeling the behavior of STDs in the sexually active community. To analyze the HIV transmission dynamics, the sexual activity and needle-sharing drug use of the susceptible population must first be understood and modeled. This is a formidable research question in itself.

Traditional models for sexually transmitted diseases [Heathcote and Yorke, Anderson, et. al.] assume that the sexually active population is homogeneously

mixed, with partners picked at random from the pool of available partners, and with the same risk per partner. However, this is a poor model for human behavior, which only works when a disease is near equilibrium. A large portion of the population has a series of partnerships, each of which lasts for a long period of time, during which there are many contacts with the same partner. More random contacts tend to occur in the periods between partnerships. Partnerships are chosen in a biased manner, from the group which is similar in age and sexual behavior [John Gagnon, personal communication:]. We are developing a model which will deal in a reasonably simple manner with both biasing in partner choice and multiple contacts with the same partner, while still allowing for interactions between high activity and low activity groups.

For example, the married man who has an affair with a married woman has a different risk from one who picks up a prostitute once a year. They may both have the same number of new partners each year, but have chosen those partners in a very different manner. Risk also depends on the infectiousness of each contact (infectiousness depends on where the infected person is in the course of the infection, the type of contact and the use of protective measures). It is perhaps important to note that the infectiousness of HIV is sufficiently low that the spouse of an infected person may not become infected until about a year before AIDS develops, so that a person's risk is not equivalent to his own risk plus his long-term partner's risk.

Endemicity of the infection also plays a major role. Once the infection becomes endemic in a group of people, it may spread in that group fairly rapidly, while another group which has few contacts with infected groups may remain protected for a long time. Age differences, physical distance, ethnicity and other social groupings may all provide barriers to the spreading of infection. Behaviors also vary between different groups of people, leading to different spreading rates in different groups.

Our model deals with all of these questions by treating each risk factor as a variable, and distributing the population appropriately. Contact rates between groups with different variable values are then estimated as well as data permits.

Because of the long time scales involved in the transmission dynamics of AIDS, members are not frozen into a given risk group once they have entered it. This flow occurs because people move from one region to another and because behavior changes for a number of reasons, including age, marital status, knowledge of infection, changing social mores and educational efforts. This is an additional

source of contact between risk groups which the models of Anderson, et al., and of Hethcote and Yorke have not dealt with. We intend to investigate the magnitude of this flow and the effect that it has on the spreading of the HIV infection.

The risk group divisions which we have identified as being of possible importance to the spreading of this epidemic area are

- a - age
- r - number of new male partners per year
- s - number of new female partners per year
- g - sexual activity group
- d - number of needles shared per year
- p - population density
- z - zone of the country
- e - ethnicity or social group

Some of these variables (a,d,p,z,e) are chosen because they act as barriers to the spread of the disease. That is, people with similar ages, ethnicity, drug use and living in nearby geographic regions are more likely to spread the virus among themselves than they are to other groups. Other variables (r,s,g,d) are crucial to accurately predicting the rate at which the virus is spreading. The time variables t and τ = time since infection are needed to follow the course of the disease.

AGE

Age is important for a number of reasons.

- The number of children born with the HIV infection will depend on the number of infected women who are having children, which varies with age.
- There is a distribution of ages at which people become sexually active and presumably a tendency to migrate first into more active groups and then into long term relationships as one ages.
- Drug use is age dependent.
- Social groups, such as college, are age dependent.
- The amount and type of traveling done also is age-dependent.
- There are natural barriers to contacts between age groups, so that the infection will not necessarily spread into all age groups in an infected area.
- Health and death rates are age dependent.

SEXUAL ACTIVITY

Risk from sexual activity depends on the probability of choosing an infected partner as well as the number and type of contacts with an infected partner. A small core of HIV infected very sexually active people (e.g. prostitutes) can drive the epidemic. The probability of choosing an infected partner depends not only on how many new partners are chosen, but also on the manner in which those partners are chosen.

Most models for the transmission of venereal diseases [Heathcote and Yorke, Anderson, et al.] have assumed that all partners are picked at random from the pool of available partners. This leads to the proportionate mixing assumption that the probability of someone with i partners per year picking an infected partner with j partners per year is i/j · (number of infected people with j partners per year)/(total number of partners picked per year). These models also assume that the number of contacts per person is the same. However, it is clear that these are overly simplistic assumptions, even for a purely homosexual model.

There is, first of all, a tendency for people with fewer partners to have more contacts per partner than people with many partners. There is also a bias of like toward like, so that people with few partners tend to pick people who also have few partners. Simply adding these biases into the Anderson, et. al., model leads to substantially different predictions (see Figures 1-3) from their random mixing model.

Another aspect of behavior is that most sexually active people, both homosexual and heterosexual, move in and out of stable partnerships [conversation with J. Gagnon, German data]. They may go into the dating pool and have a number of short term relationships, with a small number of contacts per person, before forming a new partnership, or they may go directly from one partnership to the next (with or without some overlap). The duration of the longer term partnerships tends to increase with age. A recent model for the spread of AIDS by Klaus Dietz incorporates some of these flow ideas using survey data of the West Germany population.

There is also a fraction of the population which maintains long-term relationships, and then has a certain number of outside partnerships. The risk to individuals from longer-term relationships depends on the outside partners or the previous partners of their mates. In order to deal with these ideas, we have included the ability to break up the sexual activity groups into behavior classes.

Some possible behavior classes are shown in Figure 4. We are contacting sociologists and demographers to ascertain what is known about sexual partner choice, and how many people fall into various activity and behavioral classes.

No large-scale studies aimed at ascertaining the amounts and types of sexual contact or the patterns of partner switching in the U.S. population have been conducted since the Kinsey study in 1948. This makes it difficult at this point to divide the population into behavioral groups. However, information from other studies such as fertility studies of women, and small scale studies of various group may possibly be pieced together to gain a rough picture of these demographics. For example, a survey of men in the age group 25-54 in San Francisco yielded the graphs shown in Fig. 5 on numbers of homosexual partners and proportion purporting to be homosexual, as well as some estimates of the prevalence of anal sex [Winklestein, et. al.].

Although the data is poor at this point it may be found that the infectiousness of a contact also may depend on the type of contact (male-male, female-male, male-female, anal-genital, oral-genital). Although approximately the same number of men and women are infected with HIV in central Africa, this does not imply that the virus is transmitted with equal efficiency between men and women. In central Africa the average infected male has had 32 partners in the past year (NW), some with very active females (prostitutes), while the uninfected males have had an average of only three. The women have had on the average, far fewer partners.

Infectiousness certainly depends on the use of protective devices (condoms, nonoxynol 9). We need estimates of how frequently protective devices are used, and how much behavior can be influenced by education, knowledge that a partner or oneself is infected, and by fear of infection. As public awareness increases and more people know they are infected, we speculate that the resulting change in using safer sexual practices will slow the spread of the virus. Also, individuals with higher risk behavior are more likely to seek testing and discover their infection than are those involved only in low risk behavior. This correlation between high risk behavior. This correlation between high risk behavior and being tested is easily included in the model.

In mid 1981 the first news stories on AIDS had an impact on sexual behavior in the homosexual community. The change in homosexual behavior through fewer contacts on safer sexual practices is reflected by the drop in rectal gonorrhoea in San Francisco (Pickenon Ref. 28 and 29). This happened again in 1984-85. (Does Hepatitis B show a similar trend?) Estimating the effects of future behavior

changes, although readily accommodated by the model, will be extremely important in determining the future course of the epidemic. The behavior changes could, for example, be tied to the number of active AIDS cases.

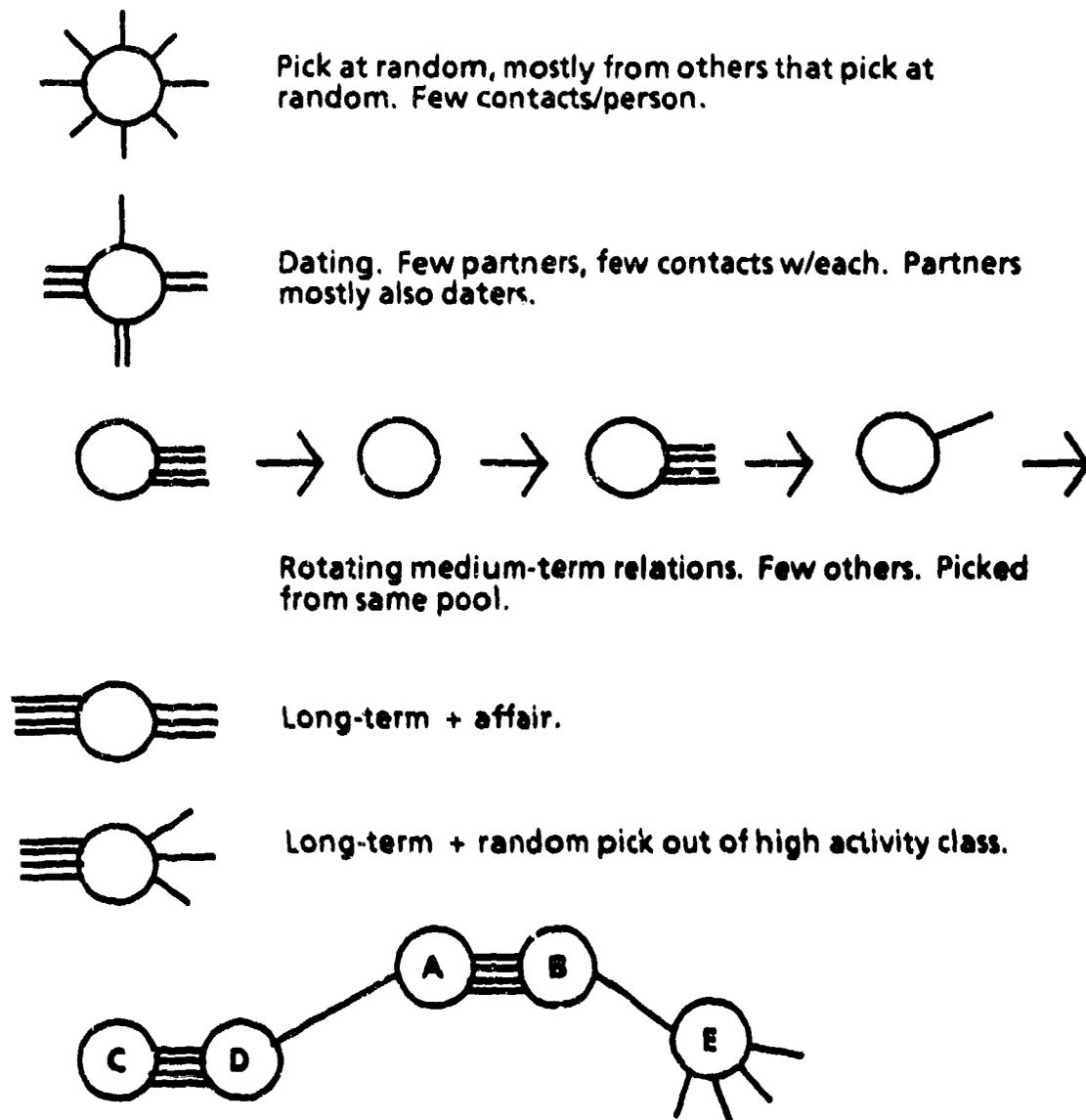


Figure 4. Different individuals (indicated by circles) may have very different sexual contacts (indicated by the lines).

DRUG USE

HIV is transmitted by sharing needles to inject drugs. Partly because many prostitutes are drug users, and partly because most drug users are heterosexuals, the spreading of HIV infection in the needle-sharing community is seen as a major source of HIV for the heterosexual community at large. We are attempting to ascertain what fraction of the population engages in needle-sharing in different age groups and regions, how they are distributed according to frequency of needle-sharing, and how much bias there is toward sharing repeatedly with the same people and against sharing with strangers [Gunzburg, Black, et. al.].

POPULATION DENSITY

The results from serological tests conducted by the Department of Defense on potential recruits indicate that the prevalence of HIV is highly correlated with population density [Burke, et. al.]. There are a number of reasons for this, each of which needs to be considered. Unlike many nonSTDs (e.g., measles, influenza) the rate of infection should not be strongly dependent upon the density of the host. People in large cities are less constrained than those in small towns. Endemicity also plays a role, since the virus will only be spread when it is present. Physical distance creates barriers between people, so that mixing may be more random and homogeneous in denser areas.

ZONES

Isolation provided by distance provides another barrier to the epidemic. To understand how rapidly the HIV infection will spread into different regions of the country, we must model how each region is connected to every other region by the movement of people. Behavior may also be somewhat regional. For example, it is unclear why HIV has spread so much more rapidly into the New York City drug community than into those in California.

Infection through blood transfusions caused a widespread geographic spread of the virus. In spring 1985, before stringent screening measures were applied to blood donors, 0.25% of the blood tested by the ELISA test were seropositive [P Ref. 15]. This led to widespread, but low level, HIV infections throughout the US; thus, seeding the susceptible population. Today, most of the HIV tainted blood is identified by the Elisa test and current blood transfusion infections will have a negligible effect on the course of the epidemic. In the model we account for this

historical misfortune by including a geographically widespread seeding of the virus. These low level infections will have a major impact on the future course of the epidemic.

The spread of the epidemic over a large territory depends upon the movement of the people. The flow of people between the major population centers by air, trains and interstate highways usually occurs with minimal contact with the intervening territory. We approximate the flow by a transfer matrix whose elements can be derived from statistical D Dot data.

The spread of the HIV into the regions surrounding the major population centers is more of a diffusion like process where the diffusion rate is a function of the population density.

ETHNICITY AND SOCIAL GROUP

The number of AIDS cases which have occurred, especially the cases in women and children, are disproportionately greater in the black and hispanic populations than in the rest of the population [Rogers and Williams]. The DOE data from military recruits also shows this bias [Burke, et. al.]. It is not understood why the infection has spread more rapidly into these populations. There are, however, barriers to contacts between different racial groups, so that it may largely be a question of endemicity. In other words, once the virus is introduced into a group of people, it can spread only in that group until a contact with a member of another group is encountered. If there are not enough contacts between racial groups, the virus can spread entirely in one group, without extending into another. These groups need not only be racial; any isolated group with few isolated contacts could experience an isolated spread. For example, students at the same university might form such a group. There may also be life style differences in these groups that facilitate spread, or a more virile strain of the virus may be introduced.

OTHER CONSIDERATIONS

NATURAL HISTORY OF HIV INFECTION

Studies of the long-term effect of the HIV virus on the immune system are all reaching similar conclusions: HIV causes a slow, but progressive, decline in the immune system. The rate of this decline varies from person to person, and some people appear to stay on a plateau for long periods of time. Short-term upward

fluctuations in measurements of quantities such as the T4 helper cells, are often observed, but most infected immune systems decline over the long run [Brodt, et al., Redfield and Burke, Melbye, et al.]. Autopsies of AIDS victims show that HIV also crosses the blood-brain barrier in a large percentage (around 80%) of infected persons and causes a wasting away of the brain; although it is not yet clear if this deterioration is a slow progression, or happens late in infection [Finkbeiner, et al.]. When the immune system is sufficiently compromised, or the brain sufficiently affected, symptoms appear. Initial symptoms of immune problems range from the very mild (so-called ARC, or generalized lymphadenopathy, or even just poor health) to Kaposi's sarcoma and the devastating opportunistic infections classified as AIDS. Deterioration of the brain leads to blindness and Alzheimer's-like dementia. Eventually, death follows. It is not clear what the appearance of KS has to do with HIV-stimulated immune system decline. KS may occur at any point more than one year after infection, independent of immune system breakdown. It is much more prevalent in homosexual men from New York City than in other groups. It is often not the eventual cause of death; immune system decline continues until an opportunistic infection back to death.

This picture of progressive immune system decline indicates that everyone infected eventually dies from HIV induced illness. However, both the time from infection to diagnosis of AIDS or dementia and the time from diagnosis to death are extremely variable. Adults have developed AIDS in less than two years (no adults develop AIDS in the first year [Weiss (1986)], and have remained well for more than eight years. In studies of patients for whom some estimates of date of infection can be made (such as hemophiliacs), the percentages developing AIDS in any given year after infection are still increasing, which leads to an estimate of an average time to AIDS of at least eight years. Death may occur anywhere from a few hours to more than six years after diagnosis, with an average of 12-14 months.

VARIABLE INFECTIVITY

Individuals who are carrying HIV vary in infectiousness as the course of the disease progresses. Studies of spouses of hemophiliacs and blood transfusion recipients indicate that infectiousness is quite small until late in the course of the infection as few spouses have seroconverted prior to a year before the onset of AIDS [James Goedert, personal communication, Fischl, et. al.]. However, partners have been known to convert immediately [Weiss, et. al. (1985)].

The infectivity may be related to the amount of free virus in the system of an infected individual. Studies indicate that the amount of free virus goes up in the first few weeks [Francis, et. al., Salahuddin] and then down as antibody response occurs, then remains at very low levels for years. As the immune system collapses in the year or so before AIDS develops, viral counts return to high levels [Robert Redfield, unpublished data]. This is schematically shown in Figure _____.

It is clear both from the spouse pair studies and from our numerical studies that the chances of infection from a single sexual contact must be quite low (less than about 10^{-2}) for most of the duration of infection, or else the virus would have spread much faster than it has. If the initial infectious period does exist, it is important that it be well-defined for infected individuals with many contacts, because it has a large effect on the rapid growth phases of the epidemic. This is especially the case when a disproportionately large percentage of infected people have only just become infected.

Any model which is going to predict the number of infected persons, cases of AIDS and deaths, must take into account the possible wide variability in infectivity. This ensures not only that approximately the right distribution of people develop AIDS or die, but also that infected people remain infectious for realistic lengths of times. A person who is healthy but infected for a long period of time has a much higher probability of infecting someone than a person who develops AIDS relatively early.

A person's infectivity also varies with immune function, and thus with time since infection. This variability in the probability of infecting someone with a single contact, is also taken care of in our model in terms of times to AIDS and times to death. Our model allows for variable infectivity and variable duration of infection by keeping track of the time since infection occurred.

TIME SINCE INFECTION (τ)

Because HIV produces a progressive decline in the immune system, the probability that an adult develops AIDS depends on how long he has been infected. The distribution of times between infection and clinical AIDS, t_A , is only partially known, due to the long times involved.

The time between infection with HIV and development of clinical AIDS is generally quite a number of years. No adults develop AIDS in the first year [Weiss, 1986], and in studies conducted so far, the percentage of people infected in any given year that has developed AIDS continues to increase. Estimates for the average

duration of infection now range from 5 to ten years, but there is a wide spread of times at which disease occurs. A schematic for the distribution of times from infection to clinical AIDS is shown in Figure _____. Because it is such a widely spread distribution, any model for the spreading of HIV needs to take this distribution of times into account. This creates a time lag between infection and AIDS cases.

Similarly, there is a wide distribution of times t_D between infection and death. Keeping track of time since infection allows us to use "best guess" estimates for these distributions. Another effect of the long duration of infection is the fact that as infection progresses more and more people will ascertain their seropositivity and change their behavior.

ENVIRONMENT OF THE EPIDEMIC

Because time for an infected person may take 5 to 10 years to convert to AIDS, the epidemic will not reach an equilibrium endemic state for a very long time. This is unlike many other epidemics, including measles [Dietz and Schenzi], gonorrhea [Heathcote and Yorke] and syphilis [Martini]. Before the natural equilibrium is reached, it is likely that medical advances and changes in life style will greatly effect the final equilibrium state. (The effects of life style changes by homosexuals in 1982-83 can already be observed in the reported AIDS cases.)

Our deterministic model smooths over the sporadic effects of local random features. However, unlike many other diseases, HIV infections persist (invisible and seeming dormant) in a few isolated individuals (with low sexual activity) for long times. This feature can cause sporadic local epidemics whenever the infected individual passes the virus to a highly sexually active person. In fact, in these situations the virus can spread rapidly, without warning, infecting a great many people. Because of the long time between HIV infection and AIDS, this situation can only be ascertained through virulent testing for the virus.

As time goes on, the environment of the AIDS epidemic is changing. Education programs are being launched to promote condom use, fewer sexual partners, use of nonoxynol-9, the use of sterile needles, etc. More people are being tested for antibodies to HIV and counseled on the implications of the test results. Treatments are being developed which will prolong the lives of infected persons, and perhaps lower their infectivity. A partially effective vaccine may eventually be developed. It will be possible to use our model to investigate the effects of each of these programs on the course of the epidemic. One will be able to ask questions

such as which groups should be targeted for educational programs given that there are limited funds available.

The probability of transmitting the infection in the highest sexually active group (prostitutes) may be significantly less than in the general population because of the increased use of condoms. The model has the option of varying the infectivity as a function of the number of sexual contacts.

DATA

Many of the sensitive facts, such as the distribution of the HIV infection, and transmission parameter values, such as the variability of infectiousness, will be known accurately only after years of careful study. The limited data will greatly influence the capability of the model to accurately predict the future. The limitations of the predictions can be somewhat understood by monitoring the sensitivity of the predictions as a function of the uncertainty in the data. These studies will also help ascertain which data is most crucial in obtaining reliable predictions. If different viral strains have different etiologies then some strains may eventually win out over others. For example, strains with longer incubation times, those that are more infectious, strains such as HIV-2 that are not recognized by the ELISA test or those that least reduce the health of the infected person when they are most infectious may eventually spread faster than other strains.

We have constructed a list of the type of data necessary to initialize the preliminary computer model and define the most sensitive parameters (Appendix B).

Another use of the computer model will be to generate estimates of unknown data that are consistent with all the known facts. For example, the past distribution of the HIV infected can be deduced from the current AIDS caseload. To determine the consistency of the generated data requires a formal mathematical model similar to the one we are designing. The available data can also be checked indirectly to determine its internal consistency by leaving some data out, generating estimates of the missing data based on the computer model and comparing the two data sets.

In our preliminary models, we have found that the uncertainties in the data and the parameter values cause changes in the details of the predictions but the trends and eventual outcome are the same. For example, the model's prediction of when a certain group will be 5% seropositive may vary between two and three years depending upon which parameters are used. But the qualitative prediction that it will be 5% seropositive is much less sensitive. Thus the relative effectiveness

of different control measures can be ascertained, even though the details, such as the exact timescale, may be wrong.

We are gathering data to estimate the rate of sexual contacts in the US from the Kinsey Institute, the Journal of Sex in American and surveys such as "How Do They Get It On in New York." We are planning trips to the Playboy mansion in Chicago and the Annual Penthouse Party to gather information first hand. This research will be vital to correctly predicting the sexual spread of the disease.

CLINICAL MANIFESTATIONS

The model could differentiate between the various clinical manifestations of AIDS based on different conversion probabilities. At this time we do not differentiate, but have a lumped conversion probability distribution which peaks between 7 and 8 years and assumes every infected individual eventually converts to AIDS. Since the conversion time may be longer in healthier populations or medical advances may lengthen the conversion time this, probability is a primary variable in the model. The modeling is further complicated because there is little data on the degree or variability of the infectiveness. Our current model assumes that a person infectiousness is proportional to the amount of virus in the blood, but we have few facts to support this assumption.

GENETIC VARIATION

The genetic variability of HIV DNA sequences indicate that the virus is mutating 5 to 10 times faster than an influenza virus [Gerry Meyers, Hahn (86)]. The variability is due primarily to duplications, insertions, or deletions of short segments and point mutations. The various strains may have dramatically different infectiousness resistance to vaccines or lead to different etiologies (e.g., KS vs PCP). If different viral strains have different etiology then some strains may eventually win out over others. For example, strains with longer incubation times, those that are more infectious, strains such as HIV-2 that are not recognized by the ELISA test or these that least reduce the health of the infected person when they are most infectious may eventually spread faster than other strains. Although at this time, insufficient data is available to include their effect in the model, we are not precluding the possibility in the future.

ANALYTIC FORECASTING

Traditional forecasting techniques will be used to identify changes in the course of the epidemic. For example, the AIDS cases diagnosed since January 1980 is well

approximated, as shown in Fig. _____, by

$$\text{total AIDS cases} = (t - 1980)^{5/2} + ? \pm 1\%$$

The total aids cases can also be well approximated by an exponential, where we have assumed safer sexual practices (fewer partners, condoms, spermicide) reduced the growth rate inversely as a function of time:

$$\text{total AIDS cases} = e^{\alpha t} + \text{_____} \text{ where } \alpha = \alpha_0/(t + 1)$$

Simplified analytic forecasting models such as these and the example in Appendix D can give good estimates on how many people will get AIDS next week, fair estimates for the next year but are insufficient to accurately predict the course of the epidemic three to five years from now. The only way to make reliable long term predictions is to include far more detail on the epidemiology and sexual behavior through full scale computer models.

CONSISTENCY CHECKS

Because the HIV epidemic is approximated as a perturbation on a stable, sexually active, population, the underlying model should mimic reality without the HIV infection included. Thus the balances between birth and death or between married and single adults must be correct before we can approximate the perturbation due to the epidemic.

GLOBAL PREDICTIONS

The model parameters could also be tuned to predict the spread of the virus in other regions. One change would be the use of condoms and spermicides. For example, they are used less frequently in Asia, Africa and Latin America than they are in the US. Furthermore, a 1978 WHO report estimates up to 26% of the adults in some of these regions are annually infected with gonorrhea. The current incidence of HIV infection in central Africa (up to 25% in metropolitan areas) raises serious political and social concerns.

CURRENT STATUS

We are working on the model development on four fronts. First we are talking to clinical epidemiologists and searching through the clinical literature in order to gain the best idea possible of the etiology of the HIV infection: the

infectiousness of different routes; the variation of infectivity as the infection progresses; the distribution of times after infection at which AIDS occurs and at which death occurs; how these variables depend on the general health of an individual, etc. Second, we are investigating the effects of adding various factors one or two at a time, such as variable infectivity, to the Anderson, et al. homosexual model. Third, we are writing the general computer model described above. Finally, we are talking to demographers and sociologists, as well as AIDS investigators, in order to ascertain what is known about sexual behavior and drug use.

This model is by no means relevant only to the spread of the AIDS epidemic. It could be applied equally to any disease by changing the parameters and the initial data file. Although, we are concentrating on the AIDS epidemic, we are producing a very versatile tool that could be used to understand the underlying transmission mechanisms of many STDs, including syphilis, gonorrhea and genital herpes. The model will help us identify the difference and similarities among these infections.

APPENDIX B REQUEST FOR AIDS DATA

REQUEST FOR DATA

We are developing a computer model at the Los Alamos National Laboratory to predict spreading of HIV into the population. This model considers spreading into males and females, into different age groups and into various regions of the country. This model attempts to incorporate all of the major transmission lines of the HIV. In order to make accurate predictions, we need to know

- efficiency of the various transmission mechanisms
- length of time that individuals remain infectious.
- variations in infectivity of individuals as the infection progresses
- current prevalence of HIV in populations with various risk behaviors
- how that prevalence is changing
- behavior changes which occur due to knowledge of seropositivity and of prevalence of HIV.

In order to obtain the above information, we are asking questions from the following general categories:

- Data from antibody tests.
- Data from infected persons.

Information about the patient's general health, when infected, how health has progressed since infection
 Information about the means by which infection occurred.
 Infection: possibly caused by the person
 Behavioral changes due to a knowledge of infection

This information is described on the following pages. Any such information will be used solely for the research purposes described above. No attempt to identify individuals will be made. Data will not be released to others without your consent.

Thank you very much. Your cooperation is appreciated.

Ann Stanley for

Tom Marr, Stirling Colgate, Pat Hagan and Mac Hyman

Data needed from antibody tests:

A. Data

This data is generated from antibody tests on various groups and consists of, numbers positive and numbers negative. Some of the following will be known about the individuals and groups tested. We need to know as much as possible about the numbers testing + and - according to

1. Sex
2. Age
3. Residence (be as specific as seems reasonable)
4. Date of test(s)
5. For males participating in homosexual activity in the past seven (7) years:
 - a. Numbers of partners in a given length of time (please specify) prior to the test.
 - b. For each partner, or an average with each partner, or over a given length of time (please specify which is given), how many times has the individual participated in
 1. receptive anal intercourse
 2. insertive anal intercourse
 3. other sexual activities
6. For males and females participating in heterosexual activity in the past seven (7) years:

- a. Numbers of heterosexual partners in a given length of time (please specify) prior to the test.
 - b. For each partner, or on average with each partner, or over a given length of time (please specify which is given), how many times has the individual participated in sexual intercourse?
7. Does the individual use condoms (always, usually, rarely)? Has this behavior changed since 1982?
 8. Does the individual use a spermicide containing nonoxynol-9 (always, usually, rarely)? If so, what is the percentage of nonoxynol-9 in the spermicide?
 9. Frequency of needle-sharing drug use.
 10. Recipients of blood products since 1970 and the dates and product type.
 11. Was the individual tested because of a known risk of infection? If so, specify the particular risk.

B. Reasons we need this breakdown:

We are developing a model which looks at the spreading of HIV in different regions of the country, in different age groups, and in males and females separately. At present, major transmission lines are by means of sexual contact and needle-sharing and by blood products in the past. The likelihood of sexual transmission depends on the type of contact, the use of condoms and Nonoxynol-9, and the number of repetitions of contact with an infected partner. It also depends on the probability that the partner is infected.

To make predictions on the course of the disease, we need to know how many people in each group are infected at present. We plan to use this data to check the accuracy of our HIV model by starting the model in the past and comparing the results to present values.

We need to know 11 so that we can understand any sample bias, i.e., how representative of the general population is it?

Data needed from seropositives and AIDS patients:

This will be data from case studies or from information forms filled out by seropositives, or their physicians. We need as much of the following information about each person as possible.

- I. Information about the patient's general health, when infected, how health has progressed since infection.
 - A. Data
 1. Age
 2. Date of first seropositive test

3. Date of AIDS diagnosis, if any, and
 - a. Illnesses first diagnosed
 - b. Date of death and cause (AIDS or non-AIDS)
 - c. Any other AIDS-related illnesses which were diagnosed later
4. Person's general health prior to infection
5. Estimate of the time period when infection occurred. How was this estimate arrived at?
6. Dates of examinations of the person and at each examination
 - a. T-4 cell count
 - b. Any indication of HIV-caused health problems
 - c. If viral cultures are taken, was it possible to grow virus from hem?

B. Why we need A.

The point of asking all of these questions is to determine: 1) probability of developing clinical AIDS as the time since infection increases; 2) the probability of dying as the time since a specific illness occurred increases; 3) how these probabilities are affected by age and general health; 4) how infectious the person may be at each stage since infection. All dates may be given in terms of time since first seropositive test.

II. Information about the means by which infection occurred.

A. Data needed.

1. Sex.
2. It it known for sure how infection occurred? If so, by what route?
3. Has the person shared needles in the past 10 years? If so, how often?
4. Does the person have one or more long-term sexual partners (more than one year duration)? If so, for each partner answer,
 - a. Is that partner infected?
 - b. How frequent is sexual contact?
 - c. Prior to knowledge of infection, were protective measures regularly employed (condoms or Nonoxynol-9)?

- d. Estimate how long the first partner was infected before the second one became infected. If the partnership began after the partner became infected, please note the length of time after infection that it started. How is this estimate arrived at?
 - e. Is this a homosexual or heterosexual partnership?
5. For males participating in homosexual activity in the past seven (7) years;
 - a. Numbers of partners in a given length of time (please specify) prior to the test.
 - b. For each partner, or an average with each partner, or over a given length of time (please specify which is given), how many times has the individual participated in
 1. receptive anal intercourse
 2. insertive anal intercourse
 3. other sexual activities
 6. a. Numbers of heterosexual partners in a given length of time (please specify) prior to the test.
 - b. For each partner, or on average with each partner, or over a given length of time (please specify which is given), how many times has the individual participated in sexual intercourse?
 7. Does the individual use condoms (always, usually, rarely)? Has this behavior changed since 1982?
 8. Does the individual use a spermicide containing Nonoxynol-9 (always, usually, rarely)? If so, what is the concentration of Nonoxynol-9 in the spermicide?
 9. Region(s) where the person lives or visits frequently.
 10. If this is an infant infected by its mother, how long was the mother infected prior to the infant?
 11. Has the person received any blood products since 1970. If so, list dates and any reasons that the infection occurred on a particular date.
 12. Are there any other means by which infection could have occurred?
- B. Reasons for these questions:**
1. We need to ascertain the ease by which the disease is transmitted by various routes. The partner-partner questions are designed to

help us estimate the probability of infection from sexual contact as it varies along the course of the disease. The mother-child questions will also shed light on infectiousness as it varies from the beginning of infection to death.

2. Our model will contain the major transmission lines for the virus. This involves breaking people up into behavioral groups, and estimating the numbers infected in each risk group. We will also look at interregional transfer of the virus.

III. Infections possibly caused by the person.

A. Data

1. If this is a woman, has she had any children since infected? If so, answer for each:
 - a. Date of birth relative to infection
 - b. Is the child infected?
 2. Is it known that this person has infected people through sexual contact, other than a regular partner. If so, when after infection did the viral transmission occur?
- B.** These questions are aimed at ascertaining infectiousness as it varies throughout time of infection.

IV. Behavioral changes due to knowledge of infection.

A. Data

1. If the person has a regular sex partner, did their sexual habits change once one was known to be infected?
 - a. Abstinence.
 - b. Decrease in frequency (how much?).
 - c. Increased condom use, (how much?).
 - d. Increased used of nonoxynol-9?
 - e. One person is very ill, so sex has decreased.
2. If this person had outside sex relations have the sexual practices changed?
 - a. Too ill to engage in outside sex
 - b. Abstinence
 - c. Condom use
 - d. Nonoxynol-9 use
 - e. Decrease in frequency (how much?)
3. If this is a woman, who planned to have children, has she changed plans for children?

- B. The reason for these questions is that we would like to be fairly realistic about the benefits of any testing plan. How does knowledge that a person is infected change an average person's behavior toward loved ones and toward outsiders? We can use our model to investigate how behavioral changes will affect the future course of the epidemic.

V. Identification of Provider

Names

Affiliations

Medical/Scientific Specialty

Description of Study that provided information

Related publications/preprints

VI. Suggested Sources for Additional Information

APPENDIX C

ECONOMIC MODELING OF THE NATIONAL COSTS OF THE AIDS EPIDEMIC

Investigators: S. Booth, R. Drake and L. Adcock

The objective is to provide economic impact predictions based on the output from the epidemiological models of the AIDS epidemic. Two categories of economic costs representing major burdens to society will be analyzed: medical system costs and national productivity losses.

Medical system costs of caring for AIDS patients will be analyzed using existing health care data bases as modified to fit the particular circumstances of the AIDS epidemic. Straightforward adding up of direct costs will show the cost burden of the disease over time. In addition, feedbacks will be incorporated into the model between the spreading of the disease over time and the bottlenecks and modifications to the health care system that will result from system overloads. Supporting financial institutions will be examined by the modeling to anticipate when cost burdens may generate failures in insurance systems, public health facilities, government budgets, and the medical community in general.

National productivity losses due to the disablement and death of citizens will be incorporated into the epidemiological models using several approaches. The

economic valuation of human life as applied in cost benefit analyses will be applied to the death rate output of the models as one measure of lost productivity. Internal demographic modeling will be important in this analysis to correctly evaluate the populations that are most effected by the epidemic. Another approach to national productivity losses will be explored for feasibility. That is to use an interindustry model of the country with production functions by industry and region. The production functions would include labor costs as inputs and so could trace the hardest hit economic sectors and their rippling effects on deteriorating national output.

The measurement of growing economic burdens as time advances resulting from various stages of the AIDS epidemic will provide crucial information to policy makers. Decisions about actions and funds to devote to the battle against the disease will depend heavily upon realistic analysis of the costs that the disease will impose.

Our work will enhance the relevance of the epidemiological model by integrating the results with measurement of the major economic costs to society. We will be breaking new analytical ground in applying production functions to the study of disease impact on national output. In addition the tools developed will be of direct relevance of potential DOD studies of the impact of AIDS and other diseases on our military readiness and war fighting capabilities.

The objective of this research is to provide analytical support in the field of economics to the epidemiological models of the AIDS pandemic. We will calculate the major economic costs to society over time. The economic work will evaluate the costs that will be borne by the medical system. It will also evaluate the losses in national output that will result from the death and disability of many citizens. This study of economic factors associated with the spread of the disease will provide new perspectives on cost benefit analysis applied to the field of health care. It is an important adjunct to the epidemiological models. It will provide quantitative guidance to policy makers for use in determining the level and timing of resources to devote to dealing with the forecasted spread of AIDS. The results of this work will be to provide readily understandable economic measures of the burden to society of the AIDS situation.

The research will be approached through a balanced application of data analysis and modeling. Existing medical and other cost data will be carefully reviewed for applicability to the AIDS situation. Problems with standard extrapolations will be carefully dealt with as the exponential spread of infection

invalidates normal data assumptions. The inadequacies of popularly quoted medical cost estimates per AIDS patient in the \$40,000 - \$150,000 range will be overcome with sufficiently detailed modeling. We will use demographic and health cost facts to provide estimates for different classes of patients and changing cost levels as numbers of cases rise. Home care, counseling, and family support costs will be considered in addition to the direct medical expenses. This economic cost submodel will interact with the epidemiological model so that linkage keeps the cost figures accurate as the model shows disease spreading at different rates through different risk groups, different regions, and other differing population criteria.

Ancillary analysis of the medical care cost results will be carried out to illuminate some of the institutional problems associated with the pandemic. At some high level of cost it can be expected that the medical insurance system will either fail or be greatly modified. The same can be said for various public health services, government health budgets, and even the private health care infrastructure itself. The economic cost submodel should be able to help to identify critical turning points in institutional susceptibility to breakdown.

The medical care cost research and submodel constitute a discrete element of the project. Medical care cost plus the national AIDS research budget are generally viewed as the societal economic burden imposed by AIDS. The second discrete element of our project deals with another economic burden which is generally ignored but is probably much larger than the health system costs. That is the productivity loss to the nation of the disabled and dead victims of the pandemic. Our approach to this part of the project involves two main analytical thrusts. One is to apply the techniques of cost benefit analyses to evaluation of the lives lost, and the other is to use an interindustry model to explore the losses resulting from work force attrition.

The most common way of calculating the economic worth of a person's life is that of discounting to the present the person's expected future earnings. This method finds wide application in cost benefit analyses used to support decisions about public works investments. In this application the analogy to public investment in AIDS research and control provides a natural and widely accepted paradigm. The method is also widely used in legal proceedings to determine damages in cases of wrongful death. The wrongful death proceedings have produced a large body of generally accepted methods for evaluating aspects of losses that can be used to individually evaluate AIDS victims, as well as, their

families. By using these methodologies in a simple economic productivity submodel, the epidemiological models' demographic information about victims can be exploited to provide robust estimates of the societal dollar losses due to AIDS deaths.

The problem with the common way of calculating the worth of a person's life is that it provides no regard for the feeling of the potential decedents. It solely concentrates on lost earnings, usually focusing on how these translate into losses to survivors. Affected individuals place a much higher value on their life than simply their expected future earnings. Following Mishan's seminal 1971 article "Evaluation of Life and Limb: A Theoretical Approach" we will try to apply some of the subsequent theoretical development to come up with more comprehensive measures of the values of victims lives. In particular, considerations of the risk borne by all potential decedents (not just direct victims) will provide another measure of the cost of the epidemic. The literature indicates that the more comprehensive measure including the potential victim's feelings will show losses several times greater than the common methodology, although with more roughness to the estimates.

Another way to measure the societal productivity losses due to AIDS is to apply an interindustry model to the situation. This is conceptually quite different from evaluating future earnings through cost benefit or wrongful death methodologies. The interindustry model uses linear algebra techniques to examine the interconnections between various economic sectors. By using production functions for the sectors, the effects of worker attrition due to AIDS can be translated into cost changes to particular industries. The industrial production functions have inputs of various types of labor, capital, and materials. Changes in the availability of labor will alter the productivity and inputs needed in most circumstances, usually in a negative way if the industry is operating efficiently before the changes. Use of an interindustry model could show very specific changes in national productivity and perhaps provide guidance to help mollify some of these effects. The application of this modeling to a pandemic would be path breaking research and involve at least two or three person-years of effort. As part of the current project, we propose simply to do some background work and planning to give a firmer idea about how to proceed with an interindustry model in the future if funding becomes available.

EXPECTED RESULTS

It is expected that during the first year we will have two economic submodels linked to the epidemiological model. The medical care cost submodel will provide the best estimates available as to the health care costs of the pandemic. Popular press estimates such as \$14 billion in 1991 reported in the February 16, TIME will be put on a robust basis and available for all of the scenarios of the epidemiological models. Similarly, the productivity submodel will provide yearly output on the losses to the economy of the dead and disabled victims of AIDS. These losses will probably be at least an order of magnitude larger than the direct medical costs. Also some preliminary analysis of health care system critical failure points will be estimated based on escalating levels of costs overtime.

Continuing work will involve improving both data sources and modeling. As the work proceeds we expect find needs for additional types of data to more realistically assess the societal impact of the pandemic. Planning will proceed on how to apply an interindustry model to productivity losses due to AIDS deaths throughout many economic sectors. Close contact with the epidemiological modelers to provide flexible response will ensure timely economic results to support the main study. The bottom line result of this work will be to provide readily understandable economic measures of the burden to society of the AIDS pandemic.

Popular press cost estimates for annual AIDS medical bills by 1991 are in the \$20 billion range and rising exponentially into the future. Productivity losses measured by standard means would be at least an order of magnitude greater than the direct medical costs. Economic losses this big would soon start to cause a deterioration in living standards. They would also impair our strategic and military strength. We must provide realistic analyses of the AIDS situation and its effect on our national well being in support of the public policy decision process that will be forced to deal with the AIDS threat to our nation.

REFERENCES

- Anderson, R. M., May, R. M., Medley, G. F. and Johnson, A., "A Preliminary Study of the Transmission Dynamics of the Human immunodeficiency virus (HIV) the Causative Agent of AIDS," *IMA J. Math. Med. Biol.* (in press).
- Black, J. L., et. al., "Sharing of Needles Among Users of Intravenous Drugs," *Lancet* 314,1 (1986) 467-447.
- Burke, D. S., et. al., "Human Immunodeficiency Virus (HIV) Infections Among Civilian Applicants for United States Military Service," October 1985 - March 1986: Demographic Factors Associated with Seropositivity. (Unpublished report, Nov. 1986).
- Dietz, K., "The Dynamics of Spread of HIV Infection in the Heterosexual Population," (unpublished report).
- Dietz, K. and Schenzle, D., "Mathematical Models for Infectious Disease Statistics," in A Celebration of Statistics, A. C. Atkinson and S. E. Fienberg, Eds., Springer, NY (1985) 167-204.
- Fischl, M. A., et al., "Evaluation of Heterosexual partners, children, and household contacts of adults with AIDS," *JAMA* 257 (1987) 640-644.
- Francis, D. P., et. al., "Infection of Chimpanzees with Lymphadenopathy-associated Virus," *Lancet*, (Dec. 1, 1984) 1276-1277.
- Ginzburg, H. M., "Intravenous Drug Users and the Acquired Immune Deficiency Syndrome," *Pub. Heal. Rep.* 99 (1984) 206-212.
- Hethcote, H. W. and York, F. A., "Gonorrhea: Transmission Dynamics and Control," Lecture Notes in Biomathematics 56 (1984) 1-105.
- Martini, E., "Betrachtungen zur Epidemiologie der Malaria und der Syphilis," *Dermatologische Wochenschrift* 19 (1928) 640-643.
- Murray, J. D., Stanley, E. A. and Brown, D. L., "On the Spatial Spread of Rabies Among Foxes," *Proc. R. Soc. London*, B229, (1986) 111-150.
- Rogers, M. F. and Williams, W. W., "AIDS in Blacks and Hispanics: Implications for Prevention," *Issue in Science and Technology*, (Spring 1987) 89-94.
- Ross, R., "The Prevention of Malaria," Second Edition, Murray, London (1911).
- Sulahuddin, S. Z., et. al., "HTLV-III In Symptom-free Seronegative Persons," *Lancet*, (Dec. 22-29, 1984) 1418-1420.
- Weiss, S. H., et. al., "HTLV-III Infection Among Health Care Workers. Association with Needle-Stick Injuries," *JAMA* 254 (1985) 2089-2093.
- Weiss, S. H. and Biggar, R. J., "The Epidemiology of Human Retrovirus - Associated Illness," *Mt. Sinai J. Med.* 53 (1986) 579-591.

Winkelstein, W. et al., "Sexual Practices and Risk of Infection by the Human Immunodeficiency Virus," *JAMA* 257 (1987) 321-325.

Hahn, B. H., et al., "Genomic diversity of the acquired immune deficiency syndrome virus HTLV-III: Different viruses exhibit greatest divergence in their envelope genes; *Proc. Nat'l. Acad. Sci.* 82 (1985), 4813-4817.

Finkbeiner, A., Hancock, E., and Schneider, S., "AIDS, Just the facts from specialists at John Hopkins," *Johns Hopkins Mag.*, Dec. 1986, 15-27.

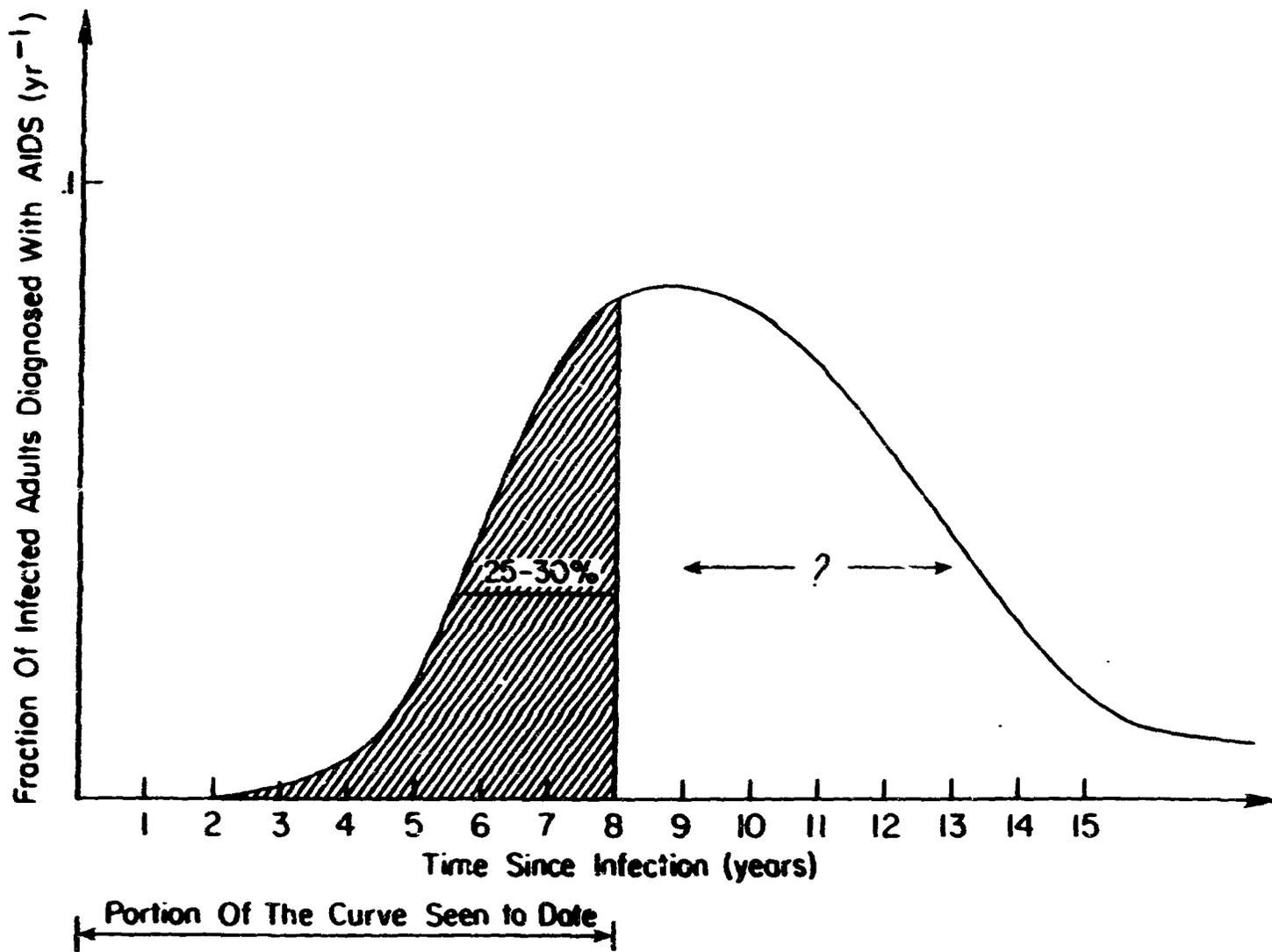
Brodts, H. R. et al., "Spontanverlauf der LAV/HTLV-III-Infektion," *Deutsche Medizinische Wochenschrift*, 111 (1986) 1175-1180.

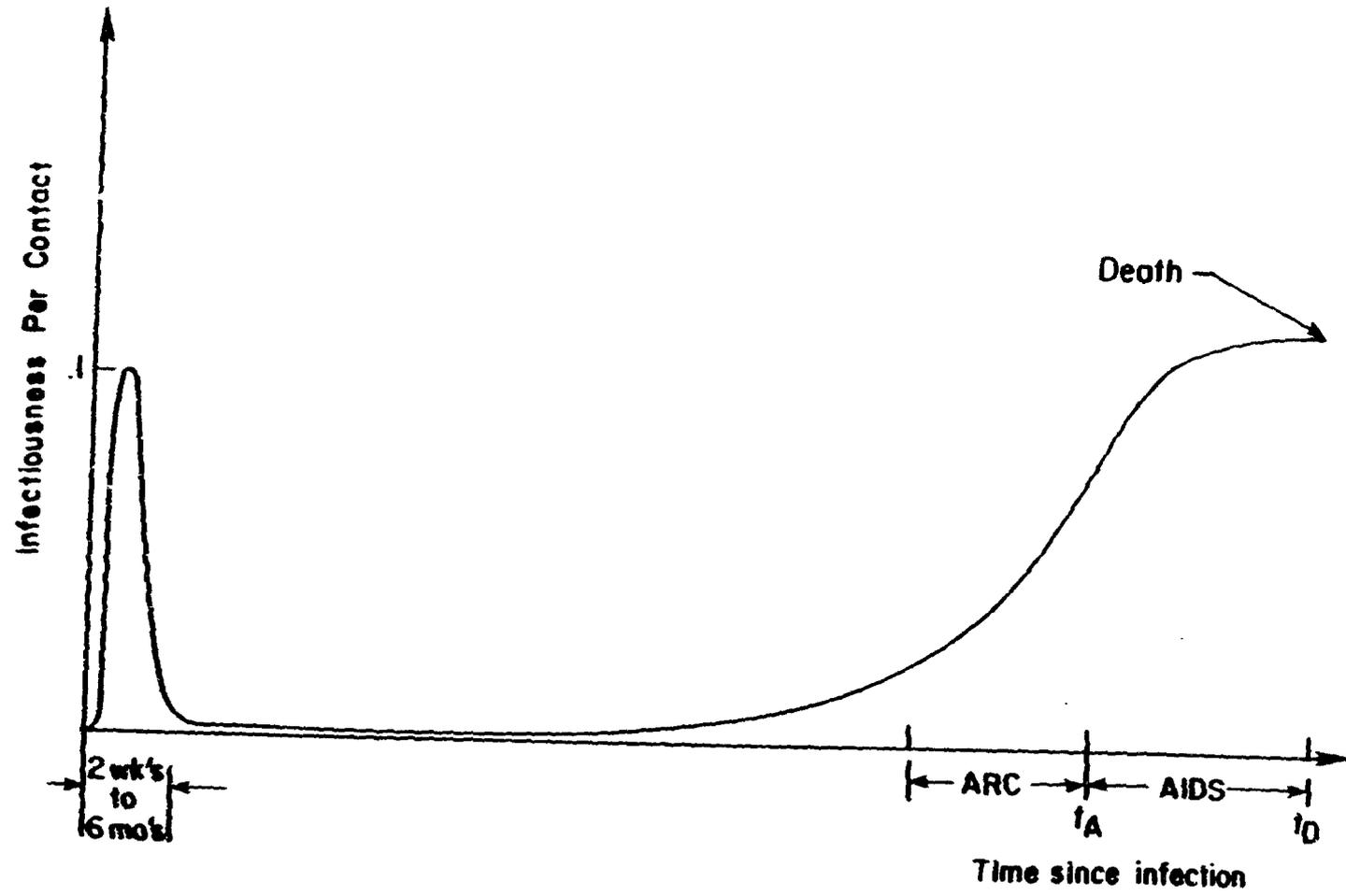
Redfield, R. R. and Burke, D. S., "Shadow on the land: the epidemiology of HIV infection," unpublished report.

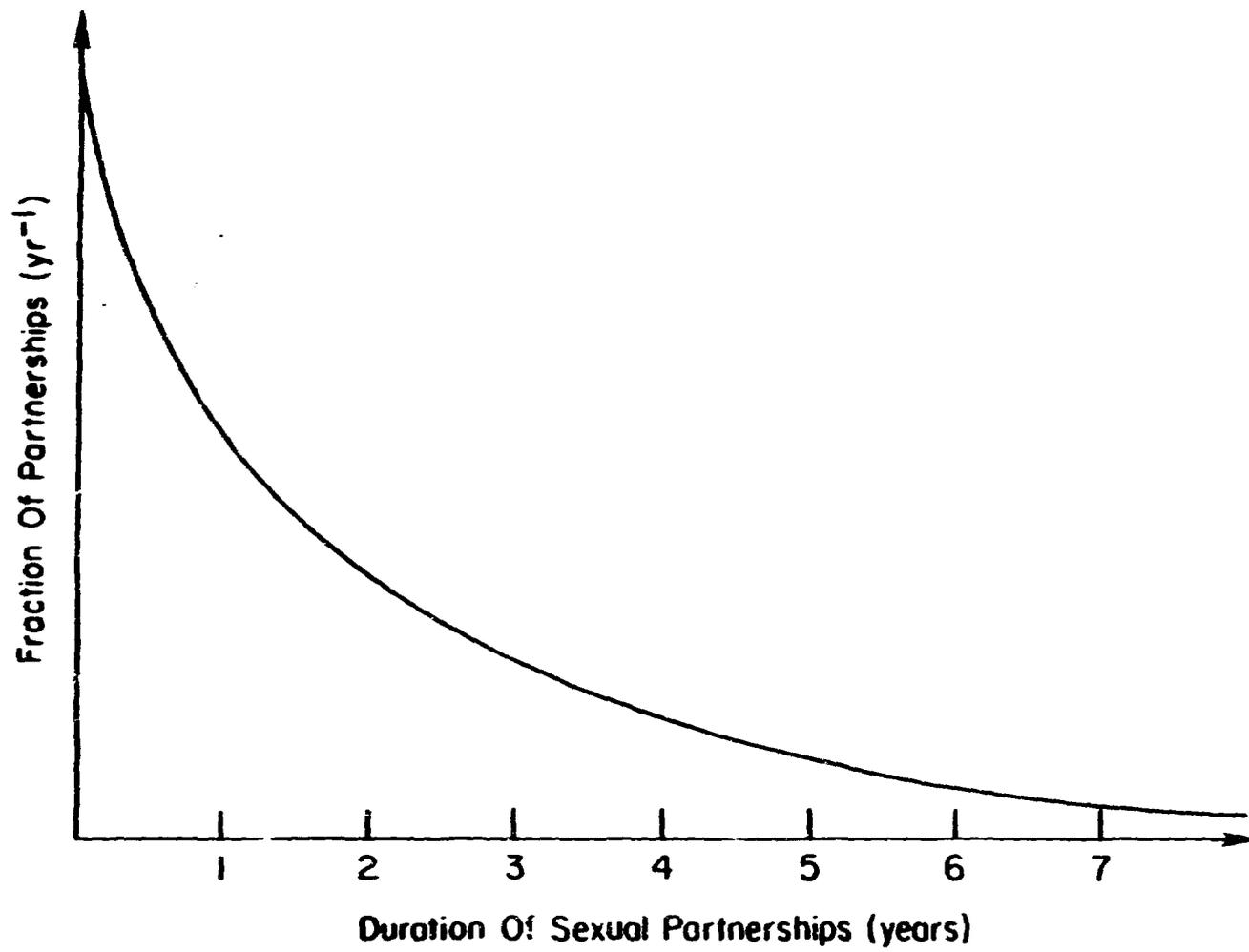
Melbye, M. et al., "Long-term seropositivity for human T-lymphotropic virus type III in homosexual men without the acquired immunodeficiency syndrome: development of immunologic and clinical abnormalities," *Ann. Int. Med.* 104 (1986), 496-500.

Lange, J. M. A., et al., "Persistent HIV antigenaemia and decline of HIV core antibodies associated with transition to AIDS," *Br. Med. J.* 293 (1986) 1459-1462.

Francis, D. P. and Chin, J., "The Prevention of Acquired Immunodeficiency Syndrome in the United States," *JAMA*, 257 (1987), 1357-1366.







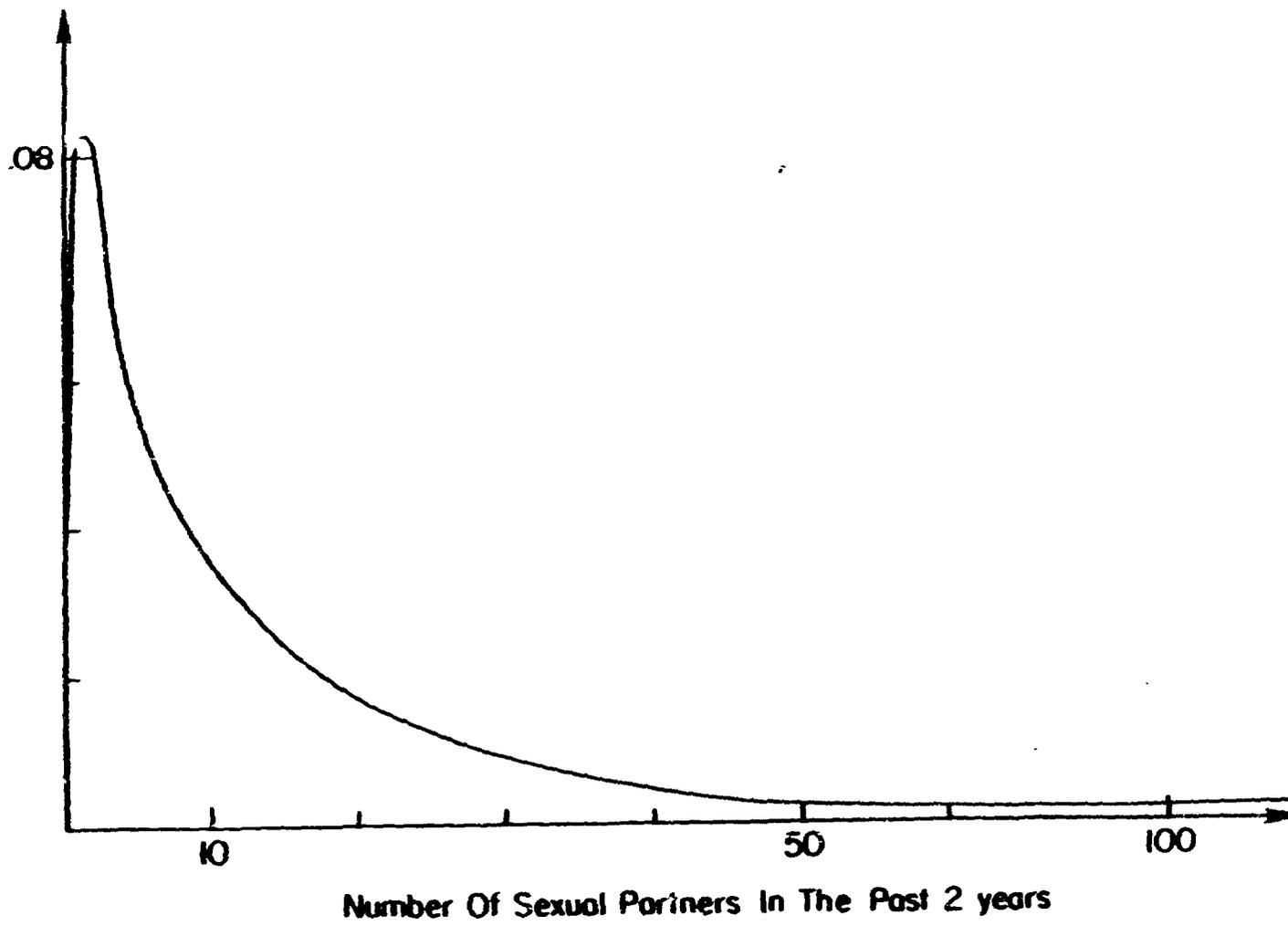


Figure 5