New England Journal of Public Policy

Volume 4
Issue 1 *Special Issue on AIDS*

Article 14

1-1-1988

AIDS: Prophecy and Present Reality

Victor De Gruttola
Harvard University

William Ira Bennett
Harvard University

AIDS: Prophecy and Present Reality

Victor De Gruttola, D.Sc.
William Ira Bennett, M.D.

Mathematical modeling of the AIDS epidemic can be useful for policymakers even though precise projections are not possible at this time. Models are useful in establishing ranges for current and future prevalence of HIV infection and incidence of AIDS, as well as in predicting the effect of a given intervention strategy. Most decision makers are using models implicitly when they use epidemiological information as a basis for policy; formulating a model explicitly permits examination of the underlying assumptions. By creating and testing a variety of models, an investigator can determine whether the models reflect more the underlying assumptions or the available data. Modeling is a process that helps the policymaker test and refine his or her own beliefs about the future of the epidemic and the effect of behavioral intervention. In this report, the process is examined in relation to five policy problems posed by the AIDS epidemic.

Because AIDS is, in many respects, an utterly novel disease, formulation of policy has proved difficult. In the normal course of events, policymakers rely on experience — often, to be sure, quantified experience — to reach decisions. The duration of the world's experience with AIDS amounts, for all intents and purposes, to half a decade. If sound policy is to be developed soon enough to be of any value in protecting public health, ways must be found to maximize the amount and accuracy of information that can be extracted from such a brief history.

In this instance, the strategy for wringing wisdom from brief experience has two major components.

First, people well versed in several different analytic strategies must collaborate to create models of the disease. Such models can be used as a basis for investigating policy options. ("Policy" in this case includes not only such obvious elements as legislative or police action to interdict transmission, but targeting and content of educational and screening programs, as well as the design of research projects.) A wide variety of models

Dr. Victor De Gruttola is assistant professor of biostatistics at the Harvard School of Public Health. Dr. William Ira Bennett is editor of the Harvard Medical School Health Letter *and is lecturer on medicine at the Harvard Medical School.*

based on very different assumptions must be available for this approach to be useful. If a certain strategy is optimal in most models, for all reasonable assumptions used to construct them, then the strategy may be regarded as "robust" to the choice of model. If a strategy makes sense only for models with very specific and untestable assumptions, then the strategy is open to serious question.

Second, models must be repeatedly tested against new data from the epidemic itself to see how close they come to being true. If models are constructed, tested, and revised in a reasonably formal, iterative fashion, many errors can be corrected before they become misleading.

Like it or not, this is the only procedure available for guiding us through the epidemic, and everyone in a position to make decisions about AIDS policy uses it, though most of the time in tacit or informal ways. The hazard of informality is that people will behave as though solutions were easier than they are, or alternatively, that they will respond to the complexity of the task with inaction.

In this article, we wish to discuss the features that make models of AIDS useful and the limitations of such models which must be explicitly recognized. We will also discuss the different types of models that may be developed; some models require complex assumptions about sexual behavior and biological characteristics of HIV infection, while others may be more strictly mathematical.

Further, we will consider five policy problems and the type of information that must be obtained to approach them: (1) whether to close sites for anonymous sex, such as bathhouses; (2) whether to promote use of condoms and "safer sex" behaviors, or, as a variant of this question, whether to concentrate resources on increasing condom use or reducing needle-sharing behavior; (3) how to use AIDS tests for screening; (4) how to budget and allocate resources for the future growth of the epidemic; and (5) what information to collect to improve the accuracy of predictions about the disease and therefore the value of interventions.

The "Reality" of the Model

Suppose that we are interested in estimating the rate of growth of the epidemic by different regions and by different routes of infection throughout the United States. In other words, we want information sufficiently fine-grained to provide an estimate of the difference between rates of growth among heterosexuals living on the West Coast and heterosexuals in New York. Two factors will determine the accuracy of the estimates.

The first is bias. Our estimate may have to be based, at least partly, on the experience of a population that does not accurately represent the population in which we are most interested. For example, the information on rate of growth of the epidemic among heterosexuals in New York may or may not be of any use in estimating the rate of growth among heterosexuals on the West Coast or among intravenous drug users in New York. If it is not, then using such information in estimating rates for other populations may give biased results.

The second factor is variance. The variance of an estimate refers to the range of values in which the estimate is likely to lie; if the range is small, then the estimate is known precisely. For example, if we are interested in rate of growth of heterosexually acquired AIDS on the West Coast, we could simply report the observed annual increase in number of cases. Such observed rates of growth are likely to be very unstable estimates of the real rate because of the small number of people who have acquired AIDS from heterosexual contact on the West Coast. Our estimate is unbiased (provided that reporting is accurate), because it is based exclusively on the population of interest; yet, it is not precise. There are so few of these cases that minor, random fluctuations in the epidemic or in reporting would have a large effect on estimates of growth rate; the variance is very high. Thus, the effect of splitting the data into fine categories is to generate values with minimal bias but high variance.

It is a fundamental limitation of model-based estimates, however, that they do not work equally well to minimize bias and to estimate variance.

An alternative approach to "splitting" is "lumping" — combining data on heterosexuals in all regions, for example, or combining data on all risk groups from one region. The number of cases is now larger, so variance is decreased. But your estimate for the specific group of interest (West Coast heterosexuals) may be highly biased if the epidemic is growing at different rates in different regions or as a function of the route of exposure to the AIDS virus.

The basic choice between lumping or splitting is itself a form of modeling. You believe either that all regions of the country are alike or that they are so different that there's nothing to be learned on the West Coast from the experience of the East Coast. Both models take extreme positions. One minimizes bias while forgetting variance; the other does the opposite. The best approach may be to construct an intermediate model that makes some trade-off between bias and variance.[1] The drawback is that this procedure requires more thought.

To estimate growth rate among West Coast heterosexuals, you could devise a weighted average combining the observed rate of growth in that specific category with an expected value of this rate based on the observed rates in other, more broadly defined categories. The weights depend on the amount of data in each category. When there are a lot of data, they are heavily weighted. But when data are sparse, more weight is given to estimates derived from related cells with larger numbers (such as all American heterosexuals or all people on the West Coast, regardless of the route of exposure to the AIDS virus). Technically, the weights are inversely related to the relative sizes of the variance estimates for the two approaches to estimation. This procedure takes advantage of the similarities between the coasts and the various groups at risk, while acknowledging that there may be important differences. The less sure we are about the differences, the more we depend on similarities to draw inference.

The point of this example is to illustrate that a model may be made more complex by having it allow for smaller and smaller effects, or by making finer divisions in region or category of risk. In general, this procedure reduces the amount of bias, but it can also reduce precision of an estimate applied to a particular region or risk group of interest, because finer distinctions necessarily reduce the number of cases in any given category.

To the extent that the obligation for AIDS planning and policy falls on states and local jurisdictions, it becomes important to recognize the perils of projecting trends in specific categories (drug users in Baltimore, heterosexuals in San Francisco) and to be aware of the techniques appropriate for attempting such projections. In particular, the trade-off between minimizing bias and minimizing variance must be kept in mind.

It is important not only to get a precise estimate of the average value of a given parameter, such as the rate of new partner acquisition, but also to estimate the variability in that rate. Even if we had an accurate estimate of the average number of new partners acquired by sexually active people but did not know the variability in this number, our models might be very misleading. There is good theoretical reason to believe that the growth of the AIDS epidemic, at least in its early phases, was driven by a small group of very promiscuous gay men. The ratio of the variance in acquisition of new partners to the mean for the gay male population at large appears, from one important study, to have been the driving factor that determined the rate of growth of AIDS.[2] This hypothesis, if true, has important implications for strategies to control the epidemic. As long as the hypothesis remains viable, it implies that collecting information about the range of behaviors is as important as getting an accurate estimate of the average number.

The point of our examples is this: The type of model you choose for analyzing any given situation is a matter of style, emphasis, and purpose. Models are designed to answer specific questions, not to reproduce reality. At the end of the 1980s, what we want from a model of AIDS is to know the small number of major factors that determine the answer to a given policy question. We want to be able to distinguish between these crucial factors and the remaining details. Nevertheless, when there is insufficient information to make a clear distinction, you may ask a simple question and still get a complex answer.

Toward a Model of AIDS

A useful model of the evolving AIDS epidemic requires many inputs and a complex sampling procedure. Models of the AIDS epidemic may be constructed for a variety of purposes. One of the most difficult challenges is to estimate the future number of AIDS cases for a given population. Currently among the most important goals of modeling is testing the potential effect of a strategy for intervention. Models can also be useful for learning what information needs to be collected, and with what precision. For example, simple modeling of the heterosexual epidemic reveals that there is little point in gathering very detailed information about heterosexual behavior unless more information about the time course of infectiousness is also gained.

The utility of models should not, however, be seen as just a slightly more quantitative way of guessing about the future than inspecting the entrails of dead birds. Models can be used to answer questions about the validity of contemporary observations (What is the evidence that behavior is really changing? Can the epidemic be propagated by heterosexual transmission? When will the epidemic peak in San Francisco?)

Information required to model some aspect of the AIDS epidemic is likely to include one or more of the following:

- the rate at which the virus is transmitted from one person to another, according to sex of partner, along with each of the potential routes of contagion (sexual activity, type of blood-borne contact, and so forth)
- how this rate changes over the course of infection
- a measure of how transmission is influenced by the presence of cofactors, such as other diseases and use of drugs
- the rate at which susceptible partners (sexual, drug using) are acquired
- the frequency and type of sexual contact between such partners

- the average duration of the relationship during which the relevant activities occur

- the extent to which the relationships and activities occur within relatively closed networks, as opposed to extending throughout the population

- the elapsed time between becoming infected and becoming ill (really the distribution of times, now known to range from a few months to more than seven years)

Because, in practice, the whole epidemic cannot be modeled, not all of these items may be needed in any one model. The point is to choose the specific question about a specific population which is to be answered, then choose the model that is appropriate for that purpose and find or estimate the value of parameters needed to construct it.

Policymakers may, and frequently do, base their decisions on their predictions of political consequences and on certain prior commitments — to values, to a constituency. These elements have weighed heavily in many policy decisions with respect to AIDS. It ought to be possible, nevertheless, to agree that the highest priority should be given to the question of effectiveness. Whether or not a particular measure violates the beliefs or behavioral preferences of various social groups, its potential effectiveness in stopping spread of the disease should weigh most heavily in judging it. Given our ignorance not only about AIDS as a disease but about the behaviors that transmit it, modeling becomes a crucial way to predict the effect of an intervention.

The following examples will illustrate aspects of this process.

Closing the Baths

Closing down places where homosexual men go for anonymous sex is one of the simpler measures that could have been taken to try to control spread of AIDS. However, the effect of such a measure depends on a rather complex interplay of factors. In this case, the effort to construct a model is itself instructive, though it may not yield an unambiguous answer about the preferable course of action (or inaction).

Closing these places may alter certain key parameters in the transmission of the AIDS virus. The rate of transmission per sexual act between an infected and an uninfected individual would probably not be affected, but the average rate of acquiring new partners might be, as would the variability in this rate, and so would the degree of mixing between infected and uninfected people.

Closing the baths could reduce the variability in the rate of new partner acquisition by making very high levels of promiscuity more difficult. However, closing centralized facilities with a certain relatively well defined clientele could have the paradoxical effect of increasing dispersal of infected people within a region. That is, it might induce former patrons to find other, less centralized outlets and thus to have a correspondingly higher chance of encountering uninfected people (who were not habitual bath-goers). On the other hand, by eliminating bathhouses as sites of sexual tourism, this measure could reduce the free mixing of people from different cities and thus slow the rate at which the disease spreads from one region to another.

A crucial parameter in any model of the effect of this action is the infectiousness of an infected person — that is, an estimate of how easily that person transmits the virus to others as a function of time since infection as well as other cofactors. How closing the baths affects the whole epidemic is influenced by the biology of AIDS and is not only a result of the effect of the measure on sexual behavior. If, for example, infectiousness is highly concentrated in a few individuals, then preventing them from having high rates of contact with susceptible people might have a major impact on the epidemic. If most infected people are infectious, but have only a low risk of infecting a partner, then closing the baths might make less difference.

For an illustration of this effect, imagine 1,000 gun owners, each of whom possesses 1,000 bullets — 999 blanks and a real one. It won't matter whether each of these people fires all 1,000 bullets at one target or at many; in any case, only one of each person's targets can wind up with a hole. Next imagine that a few gun owners have a higher proportion of real bullets, and the remainder have none. Now it makes a big difference whether the gun owners always fire at the same target or fire at many. In the second case, limiting the number of targets anyone can shoot at has a much bigger effect than in the first.

Until more information is available both about the biology of AIDS and on sexual behavior, it will be difficult to assess the effect of any intervention, such as closing the baths, aimed at limiting choice and availability of sexual partners. Any model of this effect must therefore be tested with extreme estimates of both maximal and minimal value of the parameters that are consistent with available data. "Sensitivity analysis" of this type will give a range for the likely effect of an intervention.

Of course, closing the baths, which are a social institution as well as a location for sexual activity, may have consequences that extend beyond the purely demographic. Harder to model, but nonetheless real, is the behavioral reaction of the clientele to loss of an accustomed place of sexual activity. A potential cost of closing the baths is the disappearance of a site for intensive education and behavior modification of a group that is at high behavioral risk. Alternatively, such a vivid message might be sent in the act of closing these places that behavior would be favorably influenced ipso facto.

Making Sex Safer

The object of any campaign to encourage safer sex is to alter only one parameter in the list of those causing the epidemic to grow — the infectivity per sexual act.

If the intervention being evaluated is promotion of condom use, two pieces of information are needed: the reduction in rates of transmission from wearing a condom, and the frequency of condom use.

The extent to which reducing infectivity per sexual act (as by wearing a condom) will limit growth of the epidemic depends a good deal on the variance of another parameter — promiscuity. If there is a great deal of variability in the number of partners acquired over any period of time, as appears to be the case, then targeting the most promiscuous group in the population for maximum education and intervention may have the highest yield. Even if the number of partners is not reduced, the net effect of reducing the infectiousness of behaviors among the most promiscuous may be appreciable.

It is important to bear in mind here that there is a crucial distinction between protecting a population and protecting the individual. Promoting use of condoms may prove quite effective for diminishing the rate of growth of the AIDS epidemic, especially if relatively promiscuous people use them most. But the level of protection that condom use offers any individual may, over the long run, not be very high if some of that person's partners are infected. And this is for the simple reason that risk accumulates with each repetition of an act. What may seem to be good odds on a single occasion (say, 1 in 1,000) may come to look rather poor after 10 or 50 repetitions. Thus if a program is to be judged in terms of how many participants have truly reduced their personal risk of AIDS, it must be judged by how many have really eliminated their risky behavior completely. It can be demonstrated that for a given background prevalence of infection (that is, the prevalence for the population from which partners are being selected), reducing the number of partners or using condoms does not much lower personal risk unless the number of partners is reduced nearly to one or the condoms are used nearly 100 percent of the time.[3]

On the other hand, if a program is to be judged by its overall impact on public health — how it affects growth of the epidemic — then another standard must be used. Reduction in risk that does not necessarily provide much protection for the individual might have significant effect on the epidemic as a whole.

The reason for this is that an intervention may significantly lower the basic reproductive rate of the epidemic even though it is a measure that does not assure the individual of protection. The basic reproductive rate is simply the average number of people who would be infected by an infectious person if he or she were exposed to a large group of susceptibles. This rate must be above 1 for the epidemic to propagate. If, in a given population, the rate is only slightly above 1, occasional condom use might bring it below 1. Thus, even though occasional condom use does not much alter the risk to an individual who is selecting partners from a population with a given prevalence of infection, the background prevalence itself might well be lowered if most members of the population occasionally used condoms. In any real-life situation, a model is needed to predict the magnitude of the effect.

For a simple heterosexual epidemic model in which the population is homogeneous, the basic reproductive rate can be calculated as the square root of the product of the average annual number of women infected by men (r_w) and the average annual number of men infected by women (r_m), times the duration of infectiousness (d).[4] This is expressed as $(r_w\ r_m)^{1/2}d$. If this number were close to 1, and were brought below that value by occasional condom use, then the epidemic could not propagate in that population by heterosexual contact. For example, suppose that the duration of infectiousness is 5 years, and that r_m is .1 and r_w is 1 for populations that do not use condoms and that r_m is .05 and r_w is .6 for those which do. Condom use would reduce the basic reproductive rate from about 1.6 to about .9 even though the individual using a condom with an infected partner might still be at high risk.

With respect to male homosexual practices, the impression has grown — on the basis of research reports — that anal intercourse is the most hazardous sexual practice and that oral intercourse is much less hazardous. Indeed, there is only one reported case of AIDS unambiguously attributable to oral sexual exposure.[5] However, it should be borne in mind that the power of available epidemiological methods to detect or measure the level of risk from oral exposure is exceedingly low. The value of efforts intended to shift erotic focus from one behavior to another depends on the real difference in risk between the two behaviors.

From even this short discussion, it should be clear that the value of safer sex campaigns, while undoubtedly real, should continue to be analyzed. Population-wide campaigns may be less valuable than highly targeted ones, and in any case there is an ethical obligation to distinguish between benefits that may come to the population as a whole and a level of protection that any one person may find inadequate for himself or herself.

Designing Screening Programs

There's an obvious reason for wanting to screen the population for exposure to the AIDS virus: one wants to know the total number of people infected and to know the prevalence of infection in specific subgroups of the population.

However, because of the intrinsic properties of any screening program, results are likely to be misleading for a population of very low prevalence (or, for that matter, for a group in which infection is highly prevalent). At one extreme, a large proportion of positive results is bound to be false, and at the other extreme, a sizable number of negative results will be false. For example, suppose that the test only rarely yields a false positive (say, 1 per 1,000 tests). If the test is used in a population in which infection is even more rare (say, 1 infection per 10,000 people), then a large majority (over 90 percent) of the positive test results will be false. This creates something of a paradox: you want a prior estimate of the prevalence of infection before you begin screening any particular population.

The reason for this is that you want to know how to screen to get the best estimate of overall prevalence (you don't want to concentrate your effort in areas of low prevalence). In addition, screening tells you only about specific groups like army recruits or mothers giving birth. Even if you do population-based surveys, you only find out about people who are sufficiently compliant to participate in the study. Thus, an independent estimate of prevalence is desirable, and one way to obtain such an estimate is from model-based approaches.

One such model can be constructed from data on the number of cases of AIDS in the population and the distribution of latencies (that is, the rate at which disease develops after infection, expressed not as an average but as the number of people in every interval). For another type of model, a small group is intensely studied, then its experience is projected to the whole population, making allowances for whatever is atypical about the group under study — necessarily an imprecise procedure.

In either case, modeling can help you decide whom to screen and can help in the interpretation of screening data after you get them. If, for example, your model of heterosexual transmission told you that a self-propagating epidemic among heterosexuals was impossible, then you would know that the rare cases of infection that you observed from low-prevalence areas would be more likely to have resulted from atypical behavior than from the start of a heterosexually transmitted epidemic.

Planning for the Future

There are at present reasonably good data, though far from perfect, on the growth of AIDS in the United States from the beginning of this decade. Not known — and undiscoverable from this information — is what the future growth of the curve will be.

What would make projection more precise is the shape of the curve of infection to date, but this is not known. It is not possible at present to reject either of two hypotheses: that rate of growth has been linear (increasing in a straight line) or that it has been sigmoidal (building to a peak rate of growth, then tapering off in a symmetrical fashion). The two shapes would result from different types of epidemics. If the epidemic is being spread mainly by a small number of highly infectious or promiscuous people, we may expect a linear shape, whereas if the epidemic is propagated by most of the already infected people with fairly homogeneous behavior, the sigmoidal curve may result. Existing data are consistent with either the sigmoidal or the linear model. Depending on which you believed, however, you could get estimates that ranged between 400,000 and 2 million cases of HIV infection in the United States, and projections for the future depend heavily on this estimate.[6]

It is crucial to recognize that a single model cannot be "confirmed" by showing that it fits the available data. This would be like saying, "The ground looks flat, so the world must be a plane." It is necessary to see how many different theories fit all available data. At present, a lot of different curves fit the observed growth of AIDS in the United States.

The point of continuing to construct and test these models is that, sooner or later, if one model can be rejected, we will have some useful information about the factors favoring spread of AIDS, and therefore some way of evaluating the potential effectiveness or efficiency of an intervention.

What We Need to Know

Every example we have given indicates the extreme importance of knowing more about the infectiousness of people carrying the AIDS virus. The frequency with which evidence of the virus can be recovered from genital secretions (the only important route of transmission besides needle sharing) is known, but finding the virus is a far cry from telling us what rate of transmission is typical for sexual exposure to that fluid. To get any useful information about infectiousness, large and carefully designed studies of sexual partners are needed.

Two crucial groups are those infected through blood transfusion and those infected through exposure to anti-hemophilia factor. In both of these groups, it is feasible to obtain a reasonably accurate estimate of the time of exposure. Many of the infected people also have had sustained sexual relations with a partner, so that the rate at which virus is transferred from one to another for a given frequency of sexual activity can be estimated. The essential data to gather from such people include (1) the time from probable exposure to evidence of infection, and (2) an account of sexual activity, including the type and frequency of specific acts.

The opportunity to study these groups, however, is about to be lost. New exposures have been reduced to a minimum by measures to protect the blood supply. Meanwhile, many significant events in the lives of the existing group (transmitting the virus, becoming ill) have already occurred.

Even if we had perfect information on everything else, we would still need to know about infectiousness or we would be unable to rule out the possibility that there is wide variation in the probability that AIDS will be transmitted.

The essential question about models used to develop policy is how different ones are affected by that policy. If a given policy seems best for models that make very different assumptions about unknown parameters and processes, then the policy is not dependent on more advanced knowledge. If the choice of policy is highly dependent on these assumptions, then the model has revealed where research must be conducted to obtain the vital information.

For example, a relatively robust conclusion is that heterosexual transmission of AIDS will be most effectively diminished by reducing the sharing of needles by drug users. It is clear from existing data that most heterosexual transmission will be from men who have acquired the virus through sharing of needles and who then pass it to women by way of sexual intercourse. This pattern will dominate the statistics for years to come, regardless of whether an independent epidemic can be sustained by heterosexuals.[7]

On the other hand, the effect of safer sex campaigns depends heavily on the nature of infectiousness. If the epidemic is being driven mostly by a few highly infectious and promiscuous people, then the effect of strategies like reducing number of partners or promoting use of condoms is different than if infectiousness is more diffusely distributed.

Although data are not available to model the AIDS epidemic with great precision, modeling can still be useful in making policy decisions. Uses of modeling in formulating policy include the following:

- estimating the precision with which current prevalence of HIV infection and future incidence of AIDS can be known
- making explicit the assumptions on which such estimates are based
- evaluating the likely impact of behavioral intervention or vaccination strategies
- evaluating the precision with which information must be collected to make estimates as accurate (but only as accurate) as necessary

Models need not be complex to be useful. But they must be sufficiently varied to determine under what conditions a given control strategy appears effective — does it work under a broad range of reasonable assumptions about the dynamics of the epidemic or only under special conditions? That is, models should indicate the degree to which the opinions of the modelers, as opposed to the underlying data themselves, determine the results. Modeling is a mechanism that enables the policymaker to test and refine the beliefs he or she already holds.

This research was supported by Grant No. 24643 from the National Institute of Allergy and Infectious Diseases.

Notes

1. Scott Zeger, "Methods for Trend Analysis of the AIDS Epidemic," paper presented at the Johns Hopkins Conference on Statistical and Mathematical Modeling of AIDS, Baltimore, Md., November 17, 1987.
2. Robert May and Roy Anderson, "Transmission Dynamics of HIV Infection," *Nature* 326 (1987): 137–142.
3. Harvey Fineberg, "Education to Prevent AIDS: Prospects and Obstacles," *Science* (in press).
4. May and Anderson, "Transmission Dynamics of HIV Infection."
5. Victor De Gruttola and Kenneth H. Mayer, "Human Immunodeficiency Virus and Oral Intercourse," *Annals of Internal Medicine* 107 (1987): 428–429.
6. Victor De Gruttola and Stephen Lagakos, "The Value of Doubling Time in Assessing the Course of the AIDS Epidemic," Technical Report No. 564z, 1987. Copies may be obtained from the Dana Farber Cancer Institute, Binney St., Boston, MA 02115.
7. Victor De Gruttola and Kenneth H. Mayer, "Assessing and Modeling Heterosexual Spread of HIV in the United States," *Reviews of Infectious Diseases* 10, no.1: 138–150.