BEHAVIORAL BASED MODELING OF THE AIDS EPIDEMIC

by Peter W. Plumley, FSA

This paper describes a model of the portion of the AIDS epidemic which is caused by sexual and IV drug activity. According to data published by the Centers for Disease Control ("CDC"), these modes of transmission account for over 95% of AIDS cases to date, and probably will account for an even higher percentage now that blood is screened for the AIDS virus before being used in blood transfusions. All figures in this paper exclude those AIDS cases arising from other than sexual or IV drug activity. By using the projections developed by this model together with additional assumptions as to underwriting rules, average amounts, claim costs, etc., it should be possible to get a reliable estimate of future AIDS-related insurance claims for reserving, rate-making, and other purposes.

Basic modeling approach

The basic concept of the model is that, because AIDS is transmitted primarily by certain definite activities, it is possible to develop a predictive model in terms of those activities. The model operates in three major phases. These phases develop:

1. The number of HIV+ persons;
2. The number of persons developing AIDS; and
3. The number of deaths from AIDS.

A detailed version of this paper under the title "Modeling the AIDS Epidemic by Analysis of Sexual and Intravenous Drug Behavior" will appear in Volume XLI of the Transactions of the Society of Actuaries.

In order to develop the number of HIV+ persons, the population is divided by risk group, and the number of new infections is calculated based on behavior and the probability of infection.

The number of persons infected in a calendar year form a cohort in the infected group. The number of persons developing AIDS is then calculated by applying the appropriate probability based on time since infection. The total number of new AIDS cases in a year is the sum of the new cases from each cohort.

A similar technique is used to calculate the number of AIDS deaths. The AIDS cases are maintained by cohort. Mortality rates are applied based on time since the onset of AIDS. The total number of deaths is the sum of the deaths from each cohort.

Specifically, new AIDS cases can be predicted if the following are known:

1. The number of HIV+ persons;
2. The number of sexual or drug contacts by these infected persons with persons not already infected;
3. The probability that the infection will be transmitted through a single sexual or IV drug contact; and
4. The rates of progression from HIV+ to AIDS.

The basic approach used in the model is one of trial and error. In brief, the model assumes that certain data which is known with at least some degree of accuracy is correct, and then applies a trial and error process to determine other historical data. It does this by projecting the number of AIDS cases and deaths through 1987 and then comparing the results with actual data provided by the CDC. Once this validation process is completed, it then is possible to project the number of AIDS cases and deaths and HIV infections to the year 2000, based on any desired assumptions of sexual and IV drug behavior. By varying the assumed sexual and IV drug behavior for future years, it is possible to show the effect of these changes on the progression of the disease.

The following data was assumed to be correct:

1. Total adult population figures for 1982 were drawn from data published by the Bureau of the Census.

2. It was assumed that this population included certain percentages of homosexuals, bisexuals, and IV drug users. These percentages were based on various reported studies and estimates.

3. Various rates of progression from HIV+ to AIDS were tested, based on published studies. The final scale used was based on a longitudinal study of San Francisco blood samples, and appeared to produce a level of current HIV infections most nearly in line with other estimates.

4. Rates of progression from AIDS to death were assumed, based on CDC data.

5. A probability of infection from engaging in each of the various sexual and drug acts with an infected partner was assumed. For vaginal intercourse, this probability was estimated based on published data. For anal intercourse and IV drug use, the probability was estimated without much data to verify it. However, the accuracy of these probabilities makes little difference in the accuracy of the figures, because to the extent that the risk factors are overstated, the frequencies of sexual activity will automatically be understated.

The approach used in the model is to assume that this data is accurate, and then make a guess at two sets of figures:

1. The average annual frequency of the various sexual and drug activities.

2. The number of HIV infections in 1981 for each category, and their distribution by year of infection.

Because of the substantial differences in the incidence of AIDS by race, all results were subdivided by race (black, white, and other). Also, all results were subdivided by homosexual, bisexual, IV drug user, and heterosexual. (Note: the term "heterosexual", as used in this paper, excludes IV drug users.)

Using the initial guesses of the above figures, coupled with the other data discussed above, it is then possible to project the number of AIDS cases and deaths by category from 1982 through 1987, using the formulae to be discussed later. Since the actual number of such cases and deaths is reported by the CDC, the projection then can be compared with the actual data. By a trial and error process, it then is possible to develop assumptions as to the number of 1981 infections and the frequency of sexual and drug acts that reasonably closely reproduce the known data on AIDS cases.

The model was designed to be used on a national basis, without distinguishing by geographic region. However, to the extent that data is available on a more localized basis, such as by state or a group of states, the model could be used to project the epidemic for any such locality.

To keep the model within practical limits, the approach used was to assume that, within each racial category, the various types of sexual and IV drug activity are not segregated by geography, social class, or other strata such as degree of promiscuity. Specifically, the following assumptions were made:

1. It was assumed that no interracial sexual activity takes place.

2. It was assumed that homosexuals engage in sexual activity only with other homosexuals.

3. Bisexual men were assumed to engage in sexual activity both with other bisexual men and with heterosexual women; however, it is assumed that their risk of infection comes only from other men, thus ignoring the much smaller risk of becoming infected from engaging in vaginal sex with an infected woman.

4. IV drug users were assumed to risk infection only from engaging in needle-sharing with other IV drug users, thus ignoring the much smaller risk of becoming infected from vaginal sex.

5. In order to avoid over-complicating the model, it was necessary to arbitrarily split those homosexual or bisexual men who were also IV drug users between those two categories, so that the IV drug user has a risk only from IV drugs, and the homosexual and bisexual men have a risk only from receptive anal intercourse.

6. For years prior to 1987, it was assumed that persons with AIDS were just as likely to engage in sexual or needle-sharing activities as were those who were HIV+ but who had not actually contracted AIDS. However, beginning in 1987, it was assumed that public education had progressed to the point that persons who actually had AIDS would no longer be having any sexual or needle-sharing activity unless their sexual or IV drug partner was fully protected from acquiring the disease. This approximation has the effect of slightly overstating the progression of the disease prior to 1987 and slightly understating the progression after 1987; however, as will be seen later the lack of stratification has the offsetting effect of overstating the progress of the disease after 1987.

7. For years prior to 1987, it was assumed that heterosexual men engage in sexual activity at random with heterosexual women, both IV drug users and others. However, for years beginning with 1987, a "discrimination factor" was added to the formula, so that the model could measure the effect of various degrees of avoidance of sexual activity with female IV drug users.

8. For years prior to 1987, it was assumed that heterosexual women engaged in sexual activity at random with heterosexual and bisexual men, including IV drug users. However, for years beginning in 1987, two "discrimination factors" were added to the formula. One of these permitted the model to measure the effect of various degrees of avoidance of sexual activity with male IV drug users, and the other allowed the model to measure the effect of various degrees of avoidance of sexual activity with bisexual men.

9. For years prior to 1987, it was assumed that persons generally did not have knowledge of their HIV status, and therefore that, in choosing a sexual or IV drug partner, one would not be able to exclude anyone because he or she was infected. However, for years beginning in 1987, a "knowledge factor" was added to the formulae for all categories, to permit the model to reflect the availability of information on HIV status.

Description of formulae

As previously discussed, the model assumes that the three major population groups of white, black, and other are not involved in any drug or sexual activities with each other. An initial population of N_m males and N_f females was assumed to exist in 1982, based on census data. The initial white, black, and other populations were then determined by applying factors of 0.814, 0.106, and 0.080 respectively to determine the racial split. These factors also were based on census data.

The next step was to calculate the number of homosexuals, bisexuals, and IV drug users in 1982. Based on studies of such data, an assumed percentage of 2.0% was used for male homosexuals and 1.0% for male bisexuals. These percentages were used for all three of the racial categories. For IV drug users, however, differing percentages were used by race. These percentages were as follows:

Race	Sex	Percent
White	Male	0.2%
White	Female	0.1
Black	Male	3.5
Black	Female	1.0
Other	Male	3.0
Other	Female	0.7

Those who did not fall into one of the above categories were classified as heterosexuals.

For 1982, the number of persons in each category was determined by multiplying the total population (split by race) by the various percentages shown above.

For 1983 and 1984, the number of persons in each category was determined by multiplying the figures for the prior year by 1.0135, which was estimated to be the annual population growth based on recent census data. It was not felt necessary to refine the population figures to reflect any variations by age, race or sex.

For years after 1984, however, an additional refinement was added. The projection goes to the year 2000, and involves some groups (homosexuals, bisexuals, and IV drug users) that by then will be greatly impacted by AIDS. Because of this, it was felt that for years after 1984 a downward adjustment should be made to reflect the number of AIDS deaths. Therefore, beginning in 1985 the number of persons in each category was calculated to be the arithmetic sum of the following:

(1) the population in the category for the previous year, plus

(2) the increase in population in the category from 1983 to 1984, less

(3) the deaths from AIDS in the category in the previous year.

The number of deaths from AIDS was calculated in a manner similar to that for progression from HIV+ to AIDS, by applying the progression rates from AIDS to death shown later in the paper to the number of new AIDS cases each year in each category.

The scale of progression rates from HIV+ to AIDS is very important. The more rapid the progression, the fewer the number of HIV infections for a given number of AIDS cases. The model currently is set to use the San Francisco projection rates. The resulting number of modeled HIV+ cases of 939,000 in 1988 was reasonably consistent with other estimates and studies. Table 1 shows the scale used in the projections presented here.

Table 1

Number of Years Since HIV Infection	Percent of persons who have progressed to AIDS or death: Current Year	Cumulative
1	0.3%	0.3%
2	1.2	2.5
3	2.6	5.1
4	4.3	9.4
5	6.2	15.6
6	8.6	24.2
7	10.4	34.6
8	10.0	44.6
9	8.4	53.0
10	7.2	60.2

Table 1 (cont.)

Number of Years Since HIV Infection	Percent of persons who have progressed to AIDS or death	
	Current Year	Cumulative
11	6.1	66.3
12	5.2	71.5
13	4.5	76.0
14	3.7	79.7
15	3.2	82.9
16	2.7	85.6
17	2.3	87.9
18	2.0	89.9
19	1.7	91.6
20	1.4	93.0

Rates of progression from AIDS to death were developed based on data from the CDC AIDS Weekly Surveillance Report, and are as follows:

Table 2

Number of Years Since Developing AIDS	Percent of persons dying	
	Current Year	Cumulative
0	30%	30%
1	30	60
2	15	75
3	7	82
4	5	87
5	4	91
6	3	94
7	3	97
8	2	99
9	1	100

The number of persons becoming infected each year was determined by using formulae which varied according to the particular category. For homosexuals, the formula for each racial category was:

$$VH_n = (NH_n - IH_{n-1}) \times [1 - (1 - p_a \times {}_hT_n \times (IH_{n-1} + IB_{n-1} - AH_{n-1} - AB_{n-1})/(NH_n + NB_n - (1 - {}_hT_n) \times (IH_{n-1} + IB_{n-1} - AH_{n-1} - AB_{n-1})))^H]$$

where:

VH_n = the number of new infections from homosexual activity for that racial category in year n.

NH_n = the number of homosexuals in that racial category in year n.

IH_{n-1} = the number of homosexuals in that racial category already infected by year n-1.

NB_n = the number of male bisexuals in that racial category in year n.

IB_{n-1} = the number of male bisexuals in that racial category already infected by year n-1.

p_a = the probability of infection from a single act of unprotected receptive anal intercourse.

${}_hT_n$ = the "knowledge factor" in year n for homosexuals to reflect knowledge of which homosexuals were HIV+. A knowledge factor of 1 means that there is no knowledge about who is HIV+. A knowledge factor of 0 means that everyone who is HIV+ knows that he is, and accordingly refrains from sexual activity with uninfected partners. This factor applies only to 1987 and later years.

AH_{n-1} = the number of homosexuals in that racial category with AIDS in year n-1. This factor applies only to 1987 and later years.

AB_{n-1} = the number of bisexuals in that racial category with AIDS in year n-1. This factor applies only to 1987 and later years.

H = the average number of acts of unprotected receptive anal intercourse engaged in during the year by the homosexuals in that category.

For male bisexuals, a formula similar to that for homosexuals was used, namely:

$$VB_n = (NB_n - IB_{n-1}) \times [1 - (1 - p_a \times {}_bT_n \times (IH_{n-1} + IB_{n-1} - AH_{n-1} - AB_{n-1})/(NH_n + NB_n - (1 - {}_bT_n) \times (IH_{n-1} + IB_{n-1} - AH_{n-1} - AB_{n-1})))^B]$$

where:

VB_n = the number of new infections from homosexual activity for that racial category in year n.

NB_n = the number of bisexuals in that racial category in year n.

${}_bT_n$ = the knowledge factor for bisexuals, similar to that for homosexuals described earlier, and applicable only to 1987 and later years.

IH_{n-1} = the number of homosexuals in that racial category already infected by year n-1.

B = the average number of acts of unprotected receptive anal intercourse engaged in during the year by the male bisexuals in that category.

The remaining symbols are as previously defined, and AH_{n-1} and AB_{n-1} apply only to 1987 and later years.

For male IV drug users, the formula used was:

$$VDM_n = (NDM_n - IDM_{n-1}) \times [1 - (1 - p_d \times {}_dT_n \times (ID_{n-1} - AD_{n-1})/(ND_n - (1 - {}_dT_n) \times (ID_{n-1} - AD_{n-1})))^D]$$

where:

VDM_n = the number of new infections for males in that racial category from IV drug activity in year n.

NDM_n = the number of male IV drug users in that racial category in year n.

IDM_{n-1} = the number of male IV drug users in that racial category already infected by year n-1.

${}_dT_n$ = the "knowledge factor" for IV drug users, similar to that described for homosexuals earlier. This term applies only to 1987 and later years.

AD_{n-1} = the number of IV drug users of both sexes already with AIDS in year n. This term applies only to 1987 and later years.

ND_n = the number of IV drug users of both sexes in that racial category in year n.

ID_{n-1} = the number of IV drug users of both sexes in that racial category already infected by year n-1.

p_d = the probability of infection from a single act of unprotected IV drug activity (i.e., needle sharing).

D = the average number of acts of unprotected IV drug activity engaged in during the year by IV drug users in that category.

The remaining symbols are as previously defined, and AD_{n-1} applies only to 1987 and later years.

For female IV drug users, the formula used was the female analogue of the formula for male IV drug users, namely:

$$VDF_n = (NDF_n - IDF_{n-1}) \times [1 - (1 - p_d \times {}_bT_n \times (ID_{n-1} - AD_{n-1})/(ND_n - (1 - {}_bT_{n-1}) \times (ID_{n-1} - AD_{n-1})))^D]$$

where:

VDF_n = the number of new infections for females in that racial category from IV drug activity in year n.

NDF_n = the number of female IV drug users in that racial category in year n.

IDF_{n-1} = the number of female IV drug users in that racial category already infected by year n-1.

The remaining symbols are as previously defined, and ${}_bT_n$ and AD_{n-1} apply only to 1987 and later years.

For male heterosexuals, the formula used was more complex, because it was assumed that when the male heterosexual engaged in sexual activity, his female partner might be an IV drug user, and therefore more likely to be infected. The formula was:

$$VM_n = (NM_n - IM_{n-1}) \times [1 - (1 - p_p \times {}_fT_n \times (({}_dT_n \times (IDF_{n-1} - ADF_{n-1}) + IF_{n-1} - AF_{n-1})/({}_dT_n \times NDF_n + NF_n - (1 - {}_fT_n) \times (IF_{n-1} - AF_{n-1})))^V]$$

where:

VM_n = the number of new infections for male heterosexuals in that racial category from penile-vaginal sexual activity in year n.

NM_n = the number of male drug-free heterosexuals in that racial category in year n.

${}_dT_n$ = the "discrimination factor" in year n with respect to sexual activity with IV drug users. For example, a factor of 0.8 would indicate that the average male heterosexual would be only 8/10ths as likely to have vaginal intercourse with an IV drug user as would otherwise be the case. This factor applies only to 1987 and later years.

$_{f}T_n$ = the "knowledge factor" with respect to female heterosexuals. This factor reflects the knowledge that potential sexual partners of female heterosexuals have of their HIV status. A factor of 1 indicates no knowledge; a factor of 0 indicates that all persons who are HIV+ and who might otherwisae engage in heterosexual activity with a female are aware of their status, and therefore refrain from all such activity. This factor applies only to 1987 and later years.

IF_{n-1} = the number of female drug-free heterosexuals in that racial category already infected by year n-1.

ADF_{n-1} = the number of female IV drug users already with AIDS. This factor applies only to 1987 and later years.

AF_{n-1} = the number of female drug-free heterosexuals already with AIDS. This factor applies only to 1987 and later years.

NDF_n = the number of female IV drug users in that racial category in year n.

NF_n = the number of female drug-free heterosexuals in year n.

p_p = the probability of infection for a male from a single act of unprotected penile-vaginal sexual activity.

V = the average number of acts of unprotected penile-vaginal sexual activity engaged in during the year by male heterosexuals in that racial category.

The remaining symbols are as previously defined.

For female heterosexuals, the formula becomes even more complex, for two reasons. First, it is necessary to allow for sexual contact between female heterosexuals and bisexual males, as well as IV drug users. Second, it is necessary to allow for receptive anal sex engaged in by females. The formula thus becomes:

$$VF_n = (NF_n - IF_{n-1}) \times [2 - (1 - p_v \times {}_mT_n \times ({}_bT_n \times (IB_{n-1} - AB_{n-1}) + {}_dT_n \times (IDM_{n-1} - ADM_{n-1}) + IM_{n-1} - AM_{n-1})/(({}_bT_n \times NB_n) + ({}_bT_n \times ND_n) + NM_n - (1 - {}_mT_n) \times (IM_{n-1} - AM_{n-1})))^V + (1 - p_a \times {}_mT_n \times ({}_bT_n \times (IB_{n-1} - AB_{n-1}) + {}_dT_n \times (IDM_{n-1} - ADM_{n-1}) + IM_{n-1} - AM_{n-1})/(({}_bT_n \times NB_n) + ({}_bT_n \times ND_n) + NM_n - (1 - {}_mT_n) \times (IM_{n-1} - AM_{n-1})))^A]$$

where:

VF_n = the number of new infections for female drug-free heterosexuals in that racial category from vaginal or anal sexual activity in year n.

NF_n = the number of female heterosexuals in that racial category in year n.

IF_{n-1} = the number of female heterosexuals in that racial category already infected by year n-1.

p_v = the probability of infection for a female from a single act of unprotected vaginal sexual activity.

${}_bT_n$ = the "discrimination factor" with respect to bisexual males, similar to that for IV drug users described earlier. This factor applies only to 1987 and later years.

${}_mT_n$ = the "knowledge factor" for heterosexual males, similar to that described earlier for heterosexual females. This factor applies only to 1987 and later years.

ADM_{n-1} = the number of male IV drug users with AIDS in year n-1. This factor applies only to 1987 and later years.

AM_{n-1} = the number of male drug-free heterosexuals with AIDS in year n-1. This factor applies only to 1987 and later years.

V = the average number of acts of unprotected vaginal sexual activity engaged in during the year by female heterosexuals in that racial category.

A = the average number of acts of unprotected anal sexual activity engaged in during the year by female heterosexuals in that racial category.

The remaining symbols are as previously defined, and ${}_bT_n$ applies only to 1987 and later years.

Once the number of new HIV infections was calculated, the number of new AIDS cases for each category was determined for each year by applying the rates of progression from HIV infection to AIDS, calculated as described later in this paper. The general formula used was as follows for each category (with some modification with respect to infections already present in 1981):

$$A_n = (V_{n-1} \times P_1) + (V_{n-2} \times P_2) + (V_{n-3} \times P_3) + \ldots$$

where:

A_n = the number of new AIDS cases for that category in year n.

V_{n-1}, V_{n-2}, etc. = the number of new HIV infections for that category in year n-1, n-2, etc.

P_1, P_2, etc. = the percentage of infected persons who will develop AIDS in 1, 2, etc. years after infection.

Note that no specific allowance was made for deaths from other causes between the time of HIV infection and the development of AIDS. To make such an allowance would have complicated the model in a manner not felt necessary in view of the relatively few years involved on the average between infection and the onset of AIDS, and the low average age of AIDS victims.

Overstatement of epidemic because of failure to segment

Some models have segmented the population by geographic region, degree of promiscuity, or other characteristic. As previously mentioned, this model treats all persons within a category as a single group, and does not attempt to segment them in such manner. This has the effect of overstating the number of HIV infections, because it fails to take into account the fact that to some degree or another the epidemic may be limited or contained within a subgroup.

To try to estimate the degree of this overstatement, tests were run using a simplified model based on the same general approach as the regular model, but with three segmented populations - one with a high initial level of infection, one with a medium level, and one with a low level. The results showed that so long as the proportion of persons in a subgroup that was infected was relatively low - in the 10% range or lower - combining the various subgroups into a single group had relatively little effect on the total number of cases projected. However, if the high-risk group began to become saturated, the error created by combining the groups became quite significant.

For the projections of the heterosexual cases, therefore, the failure to segment would appear to be statistically insignificant, because there is no evidence that any portion of the heterosexual community is going to have more than a very small percentage of HIV+ persons.

For homosexuals, bisexuals, and IV drug users, however, the projections presented here may well overstate the epidemic, to the extent that it will be possible to contain the epidemic within any subgroups of these categories. For example, if a large proportion of the homosexual population in one or more major cities became infected, but only a small portion of the homosexual population in other major cities was infected, the use of model on a national basis would overstate the spread of the epidemic. (It should be pointed out that the model segregates data by race, and therefore different levels of saturation by race will not create any distortion.) In the section on applications, it is suggested that the model could be used on a state or regional basis. This presumably would allow for differences in saturation by region, and reduce any overestimates.

Understatement of epidemic because of broader definition of AIDS

Late in 1987, the CDC broadened its definition of AIDS to include certain additional conditions not previously included. It has been estimated that this broader definition may increase the number of reported cases by as much as 20%. Because the assumptions in the model were validated based on the previous, more restrictive definition, the number of AIDS cases developed by the model will be less than the number reported by the CDC in 1988. In later years, this discrepancy may gradually be offset by the overstatement because of the failure of the model to segment the data, as discussed above.

Projection of the AIDS epidemic and effect of behavioral changes

At the present time, efforts are being made to persuade people to modify their sexual and IV drug habits in order to curb the epidemic. However, so far no one really knows how effective these educational efforts will be. The model can test the effect of any set of behavioral trends - it is just a matter of putting in the assumed frequencies of sexual and IV drug contacts, setting the "knowledge factors" and the "discrimination factors", and making any other special adjustments that might be desired. The model is designed to vary these assumed frequencies by individual years from 1982 through 1994, and to vary the knowledge and discrimination factors by individual years from 1987 through 1994.

If we assume a continuation of the best estimate of 1987 levels of sexual and drug activity, we get the results shown in Table 3.

Table 3

Projection of AIDS Epidemic
Assuming no Reduction from 1987 Levels of Sexual or IV Drug Activity

	Cause of HIV Infection				
Year	Homosexual	Bisexual	IV Drug User	Heterosexual Contact	Total
			Number of New HIV Infections		
1988	47,546	7,567	114,615	25,696	195,423
1989	49,423	8,055	131,264	34,739	223,481
1990	50,720	8,474	131,769	44,969	235,932
1991	51,419	8,814	115,240	55,131	230,604
1995	50,875	9,687	30,229	76,680	167,472
2000	48,294	10,492	12,553	71,209	142,528
			Number of New AIDS Cases		
1988	19,733	2,389	6,955	1,422	30,499
1989	26,270	3,299	10,686	2,184	42,439
1990	33,132	4,295	16,073	3,332	56,830
1991	38,982	5,209	23,164	4,952	72,307
1995	46,655	6,910	60,751	17,932	132,248
2000	49,111	8,215	65,354	44,590	167,270
			Number of Deaths from AIDS During Year		
1988	12,332	1,445	4,106	847	18,730
1989	17,137	2,084	6,343	1,301	26,865
1990	22,681	2,850	9,667	1,993	37,191
1991	28,386	3,673	14,308	3,006	49,374
1995	42,671	6,094	44,662	11,987	105,414
2000	48,207	7,730	65,817	35,037	156,790
			Proportion of Population Group that is HIV+		
1988	26.6%	7.2%	40.0%	0.0%	0.5%
1989	28.2	7.8	52.9	0.1	0.6
1990	29.7	8.3	65.4	0.1	0.7
1991	31.0	8.8	75.8	0.1	0.8
1995	34.7	10.2	90.8	0.2	1.0
2000	37.9	11.6	87.6	0.3	0.9

Under this projection, which may be viewed as a "worst case" scenario, the number of new AIDS cases per year will gradually increase from the 1987 level of around 21,000 to about 167,000 by the year 2000. By then, about 1.5 million persons will have contracted AIDS, and about 1.2 million will have died in the U.S. from the disease. However, it is encouraging to note that the projection shows that the annual number of new HIV infections will peak at about 236,000 in 1990, and will have decreased to about 143,000 by 2000. This indicates that the epidemic will be diminishing within a couple of years, based on the number of new infections, even under this "worst case" scenario. Unfortunately, one of the reasons for this is that the high-risk groups will have become so saturated with infections that there will be fewer of them left to infect.

The model can test any desired variations in sexual and IV drug activity. Let us make what would appear to be reasonable assumptions as to reductions in high-risk sexual and IV drug activity which will take place in the years ahead. Specifically, let us assume the following:

1. An additional 75% reduction in homosexual and bisexual activity from present levels by 1990.

2. A two-thirds reduction in IV drug activity using contaminated needles during the same period.

3. A reduction by heterosexuals in their sexual contact with bisexuals by 25% and with IV drug users by 75% over the same period.

4. Additional behavior modification arising from knowledge of the HIV status of persons with an HIV infection, as shown in the following table:

Table 4

Proportion of HIV+ persons in various categories refraining from sexual or needle-sharing activities because of knowledge that they are HIV+

Year	Homo-sexual	Bi-sexual	IV drug user	Hetero-sexual
1987	10%	10%	5%	5%
1988	15	15	5	5
1989	20	20	7	7
1990	25	25	10	10
1991	30	30	15	15
1992	35	35	20	20
1993	40	40	20	20
1994-	45	45	20	20

The author believes that with proper educational efforts the behavior modifications described above can be achieved. Therefore the projections based on these assumptions represent the author's estimate of the "most probable" projection of the AIDS epidemic, based on data available at this time. The details of these projections are shown in Table 5.

Table 5

"Most Probable" Projection of AIDS Epidemic

Year	Cause of HIV Infection: Homosexual	Bisexual	IV Drug User	Heterosexual Contact	Total
	Number of New HIV Infections				
1988	31,149	4,924	87,228	19,348	142,649
1989	20,122	3,215	71,731	17,734	112,801
1990	9,403	1,509	45,701	13,018	69,631
1991	8,421	1,353	44,545	12,975	67,294
1992	7,365	1,184	42,649	12,666	63,865
1993	6,295	1,013	41,606	12,914	61,827
1995	4,743	762	38,195	12,968	56,667
2000	2,708	432	28,928	11,900	43,967
	Number of New AIDS Cases				
1988	19,721	2,388	6,945	1,419	30,473
1989	26,175	3,284	10,563	2,154	42,176
1990	32,749	4,233	15,476	3,180	55,638
1991	37,918	5,036	21,331	4,447	68,733
1992	40,805	5,548	27,688	5,879	79,920
1993	40,973	5,677	33,898	7,342	87,890
1995	35,812	5,077	42,580	9,717	93,185
2000	19,894	2,892	42,585	11,628	77,000
	Number of Deaths from AIDS during Year				
1988	12,328	1,445	4,103	846	18,722
1989	17,106	2,079	6,303	1,291	26,779
1990	22,537	2,826	9,449	1,938	36,751
1991	27,937	3,601	13,560	2,805	47,903
1992	32,474	4,286	18,466	3,870	59,095
1993	35,480	4,777	23,836	5,077	69,170
1995	36,382	5,037	33,861	7,510	82,790
2000	25,185	3,614	42,302	10,896	81,997

Table 5 (cont.)

"Most Probable" Projection of AIDS Epidemic

Year	Number of Persons with: HIV Virus	AIDS	Cumulative Number of: AIDS Cases	AIDS Deaths
1988	939,569	36,050	84,668	48,618
1989	1,025,592	51,447	126,843	75,397
1990	1,058,473	70,335	182,482	112,147
1991	1,077,864	91,164	251,214	160,050
1992	1,082,633	111,989	331,134	219,146
1993	1,075,290	130,709	419,025	288,315
1995	1,031,148	156,050	604,362	448,312
2000	850,669	152,163	1,025,133	872,970

The projection in Table 5 indicates that the annual number of new HIV infections peaked in 1986 at about 170,000. The number already is decreasing significantly, because of safer sexual practices by homosexuals and bisexuals. Further educational efforts, combined with identification of HIV+ persons, will cause the epidemic to decrease further in terms of new HIV infections, to 62,000 in 1993 and 44,000 by the year 2000, of whom 29,000 will be IV drug users.

The number of new AIDS cases, however, will continue to increase sharply for several more years, because of the time lag between HIV infection and the development of AIDS. The peak year will be 1995, in which about 93,000 persons will develop the disease.

The number of annual AIDS deaths will similarly continue to increase, peaking in 1997 before beginning to decrease. At the peak, AIDS deaths will account for about 4% of the total deaths in the United States. This is far less than the proportion of deaths from cancer or heart disease; however, it should be remembered that AIDS deaths occur on the average at much younger ages, and that the average medical cost for an AIDS patient is very high.

The impact of the educational and identification efforts presumed in the "most probable" projection will have a significant effect on the future of the epidemic. The projection assuming no further change in sexual or IV drug behavior (Table 3) shows that the cumulative number of AIDS cases would reach 448,000 by 1993 and 1,493,000 by the year 2000. By contrast, the "most probable" projection shows that the cumulative number of AIDS cases will reach only 419,000 by 1993 and 1,025,000 by 2000, for a total reduction of 468,000 cases.

Obviously, the educational and identification efforts presumed in the "most probable" projection will be worthwhile.

Applications for the model

As has been demonstrated, the model has considerable flexibility. In addition, it is practical, in that it has enough parameters to reflect the "real world", yet few enough so that reasonable estimates of the parameters are possible. This paper has presented only two of many projections which have been made to test various possible scenarios. By being able not only to project the epidemic but to show the effect of behavioral changes on the course of the epidemic, it has a variety of useful applications. The following are some examples:

1. Projecting the epidemic on a "most probable" basis. The projection of the epidemic on a "most probable" basis should be useful for any group that needs a reliable projection for planning or policy-making purposes. Although others also are projecting the epidemic, more reliable projections should be possible by comparing or contrasting the results from the model with others using other techniques. There are many users for reliable projections of the epidemic, including government agencies, insurance companies, employers, drug companies, and health care providers.

2. Development of public policy guidelines. The ability of the model to measure the effect of behavioral changes on the epidemic is of course the most useful part of the model, and is what distinguishes it from most other approaches to modeling the epidemic. This ability should make the model useful in the development of public policy guidelines, some of which are further discussed below. It also should be of value in justifying the cost of various public policies.

3. AIDS education and prevention. The model can quantify the importance of targeting AIDS education in specific areas. It can determine the potential effect of educational programs on the epidemic, because it can show the effect of reductions in heterosexual activity, of avoidance of sexual activity with IV drug users, of avoidance of sexual activity with bisexuals, and of avoidance of sexual activity with those testing HIV+. For homosexuals and IV drug users, the effect of educational programs also can be similarly measured.

4. AIDS testing. The major benefit of AIDS testing is that those who are HIV+ will know it, and thereby refrain from spreading the AIDS virus unwittingly. Through the use of the knowledge factors in the model, it is possible to quantify the effect that testing would have on the spread of the epidemic, assuming that some portion of persons tested HIV+ would refrain from sexual or needle-sharing activities.

5. Effect of development of drugs to slow the progression from HIV+ to AIDS. The development of drugs to slow the progression from HIV+ to AIDS would have two profound effects. First, it would provide a significantly increased incentive for persons to be tested, and therefore presumably would reduce the spread of the disease by helping others avoid those who were infected. Second, by slowing the progression to AIDS, it would have the effect of increasing the number of persons who were HIV+, but had not yet developed AIDS. The model can easily quantify these effects, for any assumed level of change in the knowledge factors, once a new set of progression rates is available or is assumed.

6. Planning for health care facilities. In the absence of new drugs, the number of AIDS cases will continue to increase in the coming years, regardless of the number of new infections, because of those already HIV+. It may be desirable to develop new and

more specialized health care facilities to care for these persons. The model can not only show how many AIDS victims can be expected, but can also determine the number of HIV+ persons would need medication if a drug were shown to slow the progression from HIV+ to AIDS. These figures can be developed under any desired set of assumptions.

7. Projection of health care costs. Because the model shows not only the number of AIDS cases, but also the number of HIV+ persons and the progression from HIV+ to AIDS, it should be of value in the projection of health care costs under different assumptions as to the cost and value of various treatments.

8. Drug companies. Companies that are developing drugs to combat AIDS should also find the model helpful in determining the potential market for such drugs.

9. Projection of the epidemic by age group. Although the projections presented herein have not been segregated by age group, the model could be used to project the epidemic by age group, if the proper data were available.

10. Projection of the epidemic by state or region. At this time, the model has been run only on a national basis, and assumes no migration in or out of the U.S. However, with the proper input it can just as easily be used to project the epidemic for a particular state or region of the country, also assuming no migration in or out of the state or region. Such a projection could be useful to state authorities, providers of health care facilities, and others for planning or public policy purposes, particularly in those localities where the AIDS problem is the most serious.

11. Life and health insurance companies. The model should be of particular value to life and health insurance insurance companies because it can add reliability to long-range projections of the epidemic. Not only does the existing model do a reliable projection to the year 2000, it also is possible to extend the projection for additional years in the future relatively easily. Also, the ability of the model to quantify the effect of alternative scenarios should be useful for planning and policy making purposes. Furthermore, the ability of the model to do projections by state or region should make it of additional value for any company which operates only in certain areas of the country or has its business distributed in some unusual manner.

12. Projection of the AIDS epidemic in other countries. With proper data, the model also could be used to project the epidemic in other countries. It should be emphasized, however, that some of the key data, such as efficiency of sexual transmission, may be difficult to estimate for many countries, and may be quite different from the experience in the United States.

August, 1988