A MODEL FOR FUTURE HIV/AIDS INCIDENCE IN NEW YORK CITY

Catherine M. Crawford
Cornell University

Fred Wenstop
The Norwegian School of Management

Abstract An earlier effort at projecting the future incidence of AIDS among adult male homosexuals in New York City is reviewed in the light of data for 18 more months. One of the previous models holds up well against the new data. The model projects a temporary levelling off or reduction in new AIDS incidence, followed by a long, slower resurge. The HIV transmission probability, given an infected partner, is estimated at around 0.003. This level is barely sufficient to sustain endogenous growth in the homosexual population. From this it may tentatively be concluded that the epidemic will not start to spread exponentially among the heterosexuals provided they engage in less risky sexual behaviors than homosexuals.

INTRODUCTION

One of the greatest problems associated with the HIV/AIDS epidemic lies in our inability to accurately project how many new cases we can expect over the next few decades and into which new groups the disease will spread. One of the reasons behind this problem is that the behavioral changes necessary to stop the spread of AIDS encompass complex sociological and biological phenomenon which are not easily understood. Yet the future impact of AIDS will be a function of (1) how successful we are in understanding the association between HIV and behavior and (2) how successfully we change risky, and also quite tractable, behaviors (Crawford 1989). While exposure to the AIDS virus can occur in various modes, such as IV drug use, the transmission mode of interest in this discussion is sexual behavior. Sexual behavior in the United States has not been fully investigated and incorporated into models of AIDS primarily because information on patterns of sexual activity within a group and the mixing tendencies across groups has not been examined. Since changes in risky behaviors will influence the course of AIDS, it is important to know which sexual behaviors are risky, who is engaging in those risky sexual behaviors, and how changes in those risky behaviors can be realized and maintained (Fauci 1988).

AIDS was first discovered within the adult male homosexual population in the U.S., and the highest proportion of AIDS cases are still reported within this population. The sexual practices of adult male homosexuals, such as having multiple sexual partners and engaging in receptive anal intercourse, have been identified as behaviors which put this group at risk for AIDS (Darrow *et al.* 1987). A second population, IV drug users, is now experiencing a rapid surge in HIV incidence and prevalence and has the second highest infection rate this country. The practice of sharing needles has been identified as the behavior which puts this group at risk (Fultz *et al.* 1987). Thus, the disease has spread rapidly into two populations that have been associated with a particular type of exposure. While the heterosexual population currently has a much lower infection rate, this population could experience a dramatic rise in incidence. Accurately predicting whether this rise will occur is tricky, at best, because of the dearth of information related to sexual mixing patterns (Kaplan *et al.* 1987).

Nonetheless, mathematical simulation models of socio-biological phenomena can be employed exploratively or as a tool to estimate important values. Koopman *et al.* (1988) and Stigum *et al.* (1989) are examples of the former. They have built parameter-rich simulation models where they explore how the patterns of sexual contacts across population segments affect the spread of HIV, and show that the effects are important determinants of the propagation of the epidemic. This creates valuable understanding of the dynamics of the spread and helps indicate what kind of health policy measures would be the most efficient. Because of the lack of data and a large number of parameters, however, such models are less amenable for estimation purposes than more parsimonious models. In a previous effort to model the spread of HIV, we built a System Dynamics model that could be simulated to produce monthly numbers of new AIDS cases (Crawford *et al.* 1988). Of these, we selected the simplest one that could be calibrated to accurately reproduce the history from January 1980 to December 1986. The model introduced as few parameters as possible, but we could still choose several different sets of reasonable parameter values that would explain equally well the past and still project rather diverging futures. We concluded that it was not possible to predict long term future rates of AIDS incidence with a reasonable numerical accuracy. It was possible, however, to portray possible future modes of development, or scenarios. This would help the health authorities make contingency plans for what might be in store.

Today, we have 18 more months of AIDS data. These data have been used to review and revise the previous model so as to further restrict the range of possible scenarios. We want to estimate the HIV transmission probability and to make projections for the future incidence of AIDS cases. To make reliable estimates, the model will contain as few parameters as possible, but still retain the essential connections from the HIV transmission probability through sexual behavior characteristics and the spread of HIV to the observed pattern of AIDS development. The HIV transmission probability is estimated by making the model fit the historical pattern of incidence. A rather precise estimate of the transmission probability is achieved, but the estimate will suffer from lack of validity because of the simplifications built into the model. To partly compensate for this, the presumed impacts on the estimate by the various simplifications are discussed. Our projections of the future incidence of AIDS are not subject to this problem in the same way, since we limit ourselves to demonstrating that simple mathematical models can explain the past well, and still produce rather different futures.

THE MODEL STRUCTURE

The population consists of male homosexuals/bisexuals in New York City, excluding intravenous (IV) drug users and prison inmates.[1] The model captures the major epidemiologic and behavioral dimensions associated with HIV/AIDS.[2] Three stages are used to differentiate the population according to their HIV status. The first stage corresponds to the susceptible population in an epidemic model and includes those persons not infected with HIV/AIDS. It is assumed that the population is not sexually inactive. The second stage corresponds to the infectious population in an epidemic model and includes all persons who have been exposed to HIV and have become infected. It is assumed that no one is immune and all who become infected are immediately contagious. Also, based upon the simplifying assumption that those diagnosed with AIDS will no longer engage in unsafe activities, contagiousness ends with diagnosis. The final stage corresponds to the population which is removed from an epidemic model. Placed within this final stage are all those who have died from AIDS according to either the CDC surveillance definition AIDS or from HIV related disease.

Four factors determine the probability of a random person becoming infected. The first factor, the transmission probability, is the probability that a person will be infected by one exposure to an infected individual due to a typical risk act. The other factors are represented as the number of new sexual partners one has per month, the number of risky sex acts performed per partner, and the probability that a new partner is contagious (seroprevalence).

Simplifying Assumptions The following assumptions have been made in order to make a reasonably simple model which still captures the major impacts of the most important behavioral parameters:

(1) There are no new entrants to the population. Essentially then, this is a cohort study. The consequence is that long-term projections will be on the optimistic side.
(2) The population is divided in two subpopulations with no sexual cross-communication. The core group is more sexually active than the peripheral group. We have determined that if we exclude behavior change, a single group model is too simple to explain the historical data. A two group model is the simplest multi-group model that remains rich enough to explain the slower growth observed from 1986. Simulations show (Koopman et al. 1988) that, for a given level of the HIV transmission probability, the onset of the epidemic will be faster if we exclude contacts between the groups, than if we allow contacts. The reason is that the more sexually active group "wastes" energy infecting the other. Thus, with contacts, we would need a higher transmission probability to explain the data. This, therefore, contribute towards an underestimation of the transmission probability.
(3) Sexual partners are selected perfectly at random within each group, which is a gross simplification. In reality, there are many barriers between population segments within the groups, notably between different age levels. However, whereas barriers slow the spread, group coherence breeds the virus. The net effect may not be great.[3]

(4) Variability in sexual behavior is an important factor for the spread of the virus. By assuming the groups are internally homogenous, we have made one step towards accounting for that. It is an open question whether we would earn anything in terms of generality by making the groups more homogenous by introducing more of them.
(5) There is no change in behavior over time, i.e. no adoption of safer sexual practices or less sexual activity. This is introduced to be able to separate the effects of group segmentation and behavior change. There are indeed reports that indicate that behavior changes have been occurring (Winkelstein *et al.* 1987; Martine 1987). If this is true, our results would be more pessimistic than necessary, counteracting the effects of the cohort assumption in assumption (1) above.
(6) All people who are infected become contagious. This is probably a reasonable assumption.
(7) All infected people are equally contagious. The level of contagiousness is constant over time from the point of infection until the possible development of AIDS, when it vanishes. Evidence indicates that contagiousness declines right after infection, and then increases to a peak after several years (Crawford et al. 1988). Our use of a constant number to describe the virus transmission probability must therefore be interpreted as some sort of an average. Since the real contagiousness apparently involves a time delay, our transmission probability would be a low estimate of the real average infectivity.
(8) The probability of transmitting HIV in a risk act is independent of the previous number of risk acts. This is a very convenient and probably harmless simplification.
(9) The incubation time is gamma distributed. It is well attested (Bachetti and Moss 1989) that the incubation time distribution has a peak after 5 to 10 years, and then a possible long right tail. The gamma distribution is the simplest distribution with this form. It has only one parameter.

Model Parameters The fixed parameters include: the estimated size of the population[4] (300,000); the probability of eventually developing AIDS when one is infected (0.6, or 1.0); the number of new partners per month (average over both groups 0.8); the relative sexual activity between the two groups; the number of risk acts per partner per month (10), the average sero-conversion time (1.5 months); and the average incubation time (7.5 or 10 years). The free parameters are the number of infected persons in each group as of January 1980 and the probability of receiving the virus in one risk act with an infected partner (the virus transmission probability).

Previous Results In the previous study based upon data up to December 1986 (the data were available in August 1987), it was determined that an assumption of behavior change or an assumption of two groups with different sexual activity would be sufficient to explain the data (Crawford *et al.* 1988). Since behavior changes can be projected in many ways, we found it more instructive to consider a pure two-group model. With an average incubation time of 7.5 years and a probability of eventually developing AIDS when one is infected of 0.6, we could make an excellent fit with the inclusion of two assumptions: (1) that 2700 persons were infected in January, 1980 and (2) that the HIV transmission probability in one risk act is approximately 0.004.

Actually, with the parameter values just quoted, we still had freedom in selecting the relative sexual activity (the number of new partners per month) of the two groups. In Table I two scenarios are shown. In Scenario A, the core group is 10% of the whole population and it was assumed they had 2.6 new partners per month compared to 0.6 new partners for the peripheral group. In Scenario B the size of the core group remains constant at 10% of the total population. The levels of sexual activity are 3.5 new partners and 0.5 new partners respectively. The average sexual activity for the whole population in both cases is 0.8.

Table 1: Parameters Comparing the Impact of Differences in Sexual Activity

	Scenario A	Scenario B
Incubation Time In years:	7.5	7.5
AIDS probability:	0.6	0.6
Number of New Partners/Month		
Core Group (10% of population):	2.6	3.5
Peripheral Group:	0.6	0.5
HIV Transmission Probability:	0.0041	0.0035
Jan. 1980 HIV Level (Cases):	2700	2700

The two scenarios fit the data equally well, but there is no way to decide which one is the most accurate. The interesting point is, however, that even though we can make excellent fits to historic data with different assumptions about relative sexual activity, the differences have big impacts on projections into the future. This is shown in Figure 1, where we see that although the pattern of the projected curves are similar, the actual incidence levels are very different. The two group model projects an imminent levelling off followed by a resurge. The numerical values of the projections are very sensitive to assumptions of relative sexual activity between the two groups. The boxes represent actual data. The reason for the levelling off in the projected pattern is that the core group is becoming saturated. The resurge is due to the peripheral group where the HIV initially spread less rapidly, but starts to dominate the picture when the core group decline.

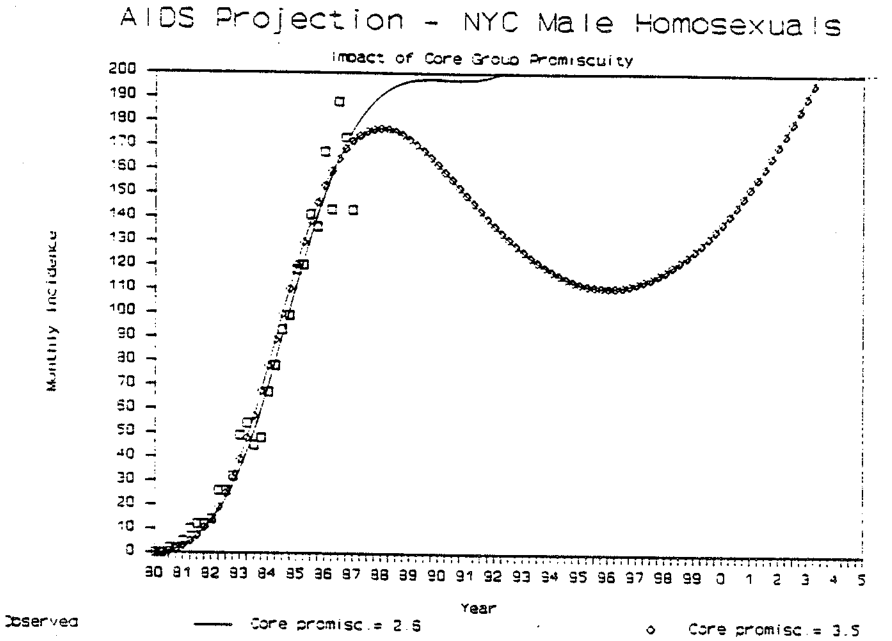

Figure 1: The Impact of Number of New Parameters on the Two Group ModelThe

concept of two groups with different levels of sexual activity is an arbitrary abstraction. It is the simplest way to model variance in sexual activity. Reality is much more complex. We concluded therefore, that this result indicates that one cannot make long term prognoses for AIDS incidence with a reasonable degree of accuracy. One can, however, make projections which portray possible futures; futures that cannot be ruled out with the current knowledge.

REVISED RESULTS

New data for AIDS incidence up to August 1988 were available in December 1988 from the New York State Health Department. They consisted of 18 new observations. At the same time, the old data have been revised and consolidated. Figure 2 shows the new data compared with the old data and a fitted curve based on the old data in Scenario B. The fitted values lie significantly higher than the new data. This is due to the revision of the old data, which mainly consists in leaving prison inmates out. Figure 3 shows how Scenario B compares with the new data. The old data estimate of the HIV transmission probability was 0.0035. The old projections apparently tend to overestimate new incidence data. This is mainly due to the revision of the data and makes it necessary to recalibrates the model.

Recalibration Based on the New Data We used Scenario B as our starting point and estimated the free parameters to make the model fit the new data. The result is that the HIV transmission probability was estimated to be 0.0030, compared to 0.0034 in Scenario B. All other parameter values are the same. We shall refer to this new scenario as C. The fitted curve is shown in Figure 4.

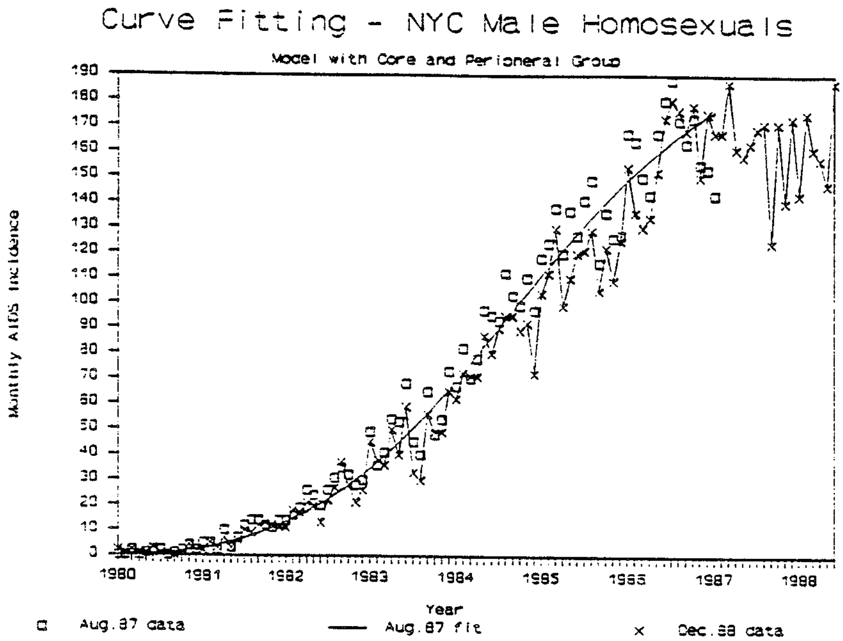

Figure 2: Comparison of Old and New Data Sets Using Scenario B

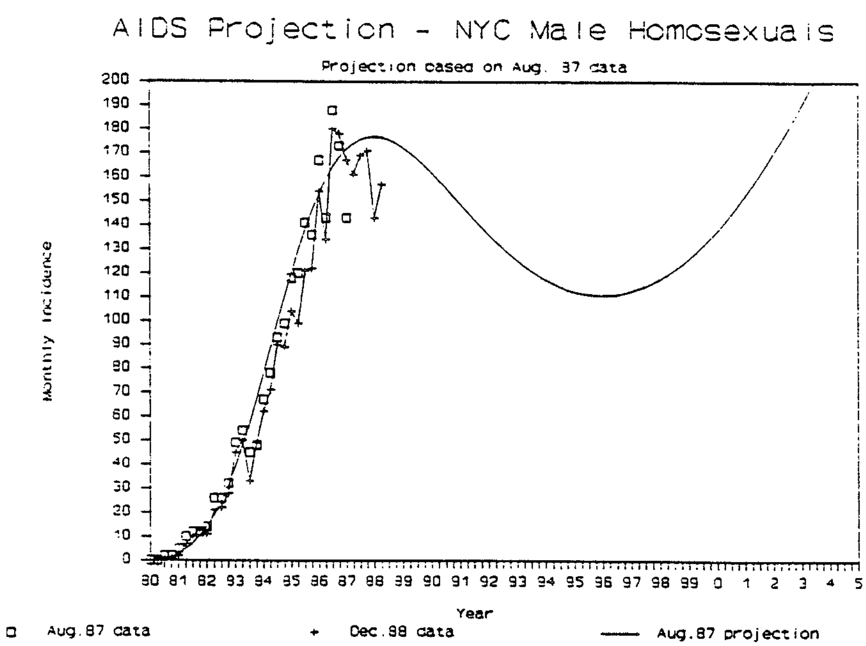

Figure 3: Comparison of the Old Projections and the New Data

Recent evidence indicates that the average incubation time may be closer to 10 years, compared to the 7.5 years we used in Scenario C. To make the model fit the new data in this case, we had to make several adjustments in the parameter values.

The result is called Scenario D. Here, we assume that the core group is 8% of the size of the total population, that the number of new partners per month is 0.35 and 6 for the peripheral and core group respectively, and that 2000 are infected as of January 1980. We estimated the HIV transmission probability to be 0.0025.

Table 2: Scenarios Based Upon New Data

	Scenario C	Scenario D
Incubation Time In years:	7.5	10.0
AIDS probability:	0.6	1.0
Number of New Partners/Month		
Core Group (8% of population):	3.5	6.0
Peripheral Group:	0.5	0.35
HIV Transmission Probability:	0.0030	0.0025
Jan. 1980 HIV Level (Cases):	2700	1400

The fitted curves derived from Scenarios C and D are compared in Figure 4. The corresponding projections of future AIDS cases are shown in Figure 5. The projection in Scenario D, based upon an incubation time of 10 years, portrays a much more

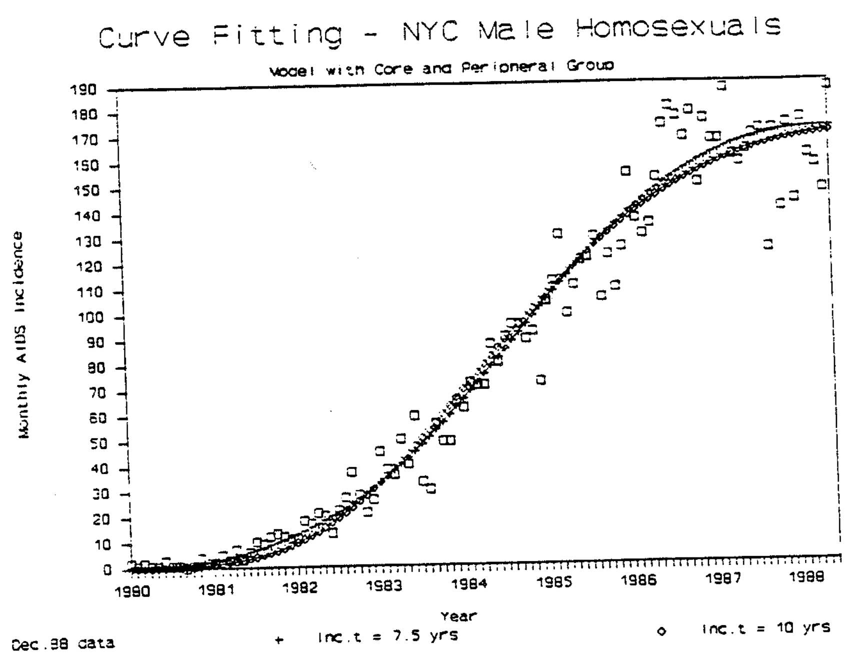

Figure 4: Model Fitted to New Data Using Scenarios C and D

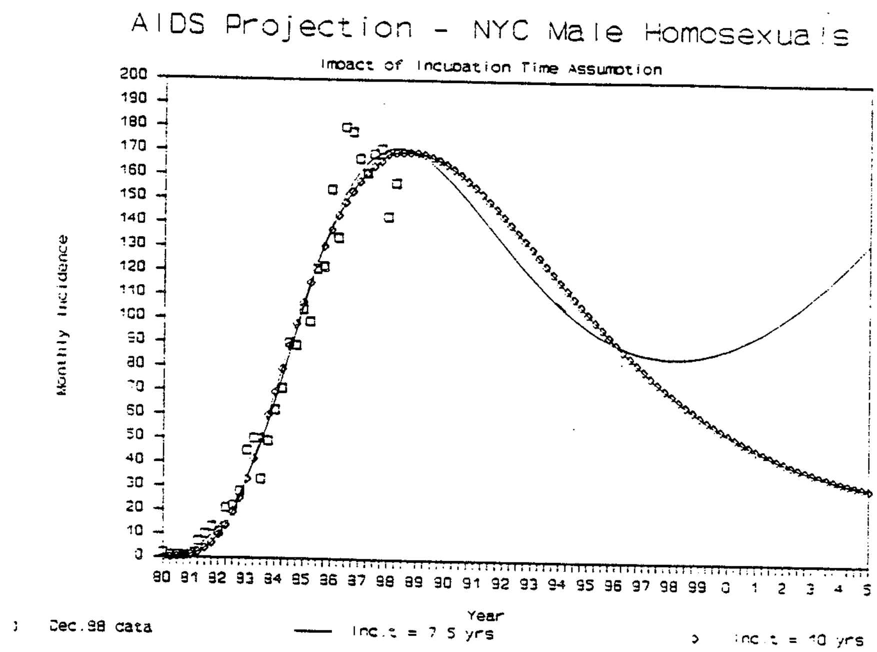

Figure 5: Projections of AIDS Incidence Based on New Data and Scenarios C and D

sluggish process than the other. The curve will eventually pick up again, outside the time horizon in the figure. It is interesting to notice the similarity in the patterns of the two projections in Figure 5, and how different the numerical values are. This is the same phenomenon we observed in the previous work. A possible future for New York City male homosexuals and bisexuals, provided there are no change in behavior, is therefore a temporary downturn in the rate of new AIDS cases, followed by a long, slow increase caused by an endogenous spread in the peripheral group.

THE CASE OF HETEROSEXUALS

The New York State incidence rate for heterosexuals averaged in 1988 around 20 new cases per month, in 1987 around 15. Such small numbers can easily be explained by contamination from other groups, and need not be due to an endogenous spread. The number of cases is rising, however, and it is expedient to ask whether an endemic spread may be just starting among heterosexual. This question will be analyzed by comparing the heterosexual situation with the homosexual case.

In the homosexual model, an average number of 0.8 new partners per month and 10 risk acts per partner was assumed. We further assumed a constant level of contagiousness after infection. Based upon this, we used the incidence data to estimate the probability of transmitting the virus to an uninfected partner in one risk act to be about 0.003. Continuing to work with the two group model for homosexuals and using the same values for sexual activity, we now ask: By how much must the HIV transmission probability decrease before the incidence rate in the peripheral group tapers off and never rises again? That would mean that the epidemic cannot spread endemically in the group.

It turns out that in Scenario C, the HIV transmission probability must be reduced by 90% to halt the spread. In Scenario D, the previously estimated value of 0.0025 for the transmission probability happens to be just on the borderline. Even a small reduction is sufficient to halt the endemic spread. We have presently more reason to believe in Scenario D than in C (Bachetti and Moss 1989). It thus appears that the heterosexual population need to engage in only slightly less risky sexual behavior than the homosexuals to avoid an endogenous spread of HIV.

The probability p_i, that a random person becomes infected during a month is determined by the number of new partners, n; the number of risk acts per partner, a; the HIV transmission probability , p_t; and the sero prevalence, f.

$$p_i = 1-(1-(1-(1-p_t)^a)f)^n$$

Of course, instead of reducing the transmission probability, which would imply an adoption of safer sexual practice, we could have reduced the number of new partners per month or the number of acts per partner. This would imply less sexual activity.

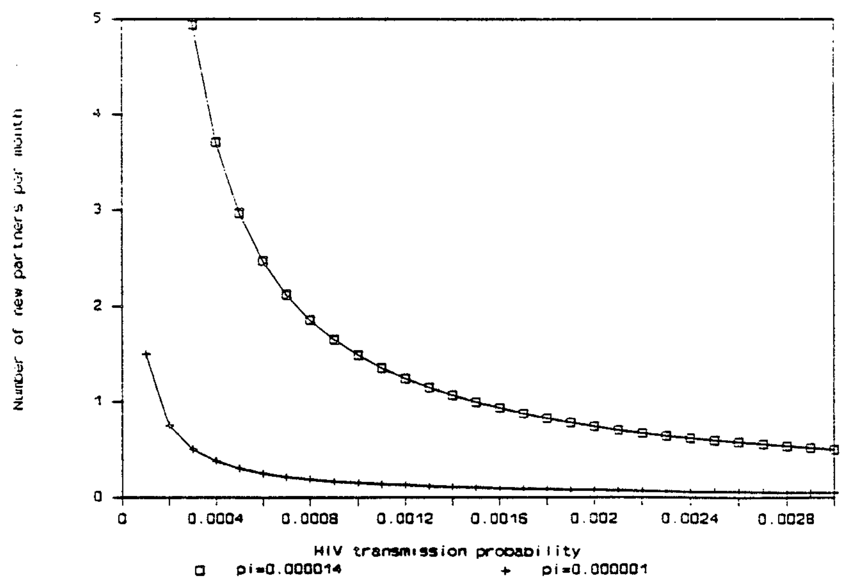

Figure 6: Different Levels of Sexual Activity Can Produce the Results in Scenario C.

It would be interesting to find out how much the level of sexual activity would have to be reduced to match the effect of an increase in safety. The variable that determines the rate of HIV spread in the model is the probability, p_i, that an uninfected person becomes infected during a given month. The formula that connects p_i with the behavioral parameters is developed in Crawford *et al.* (1988).

Figure 6 shows the trade-off between safety (p_t) and sexual activity (n) for two different levels of HIV risk (p_i). The HIV prevalence (f) is assumed to be 0.001, but its value does not alter the trade-off values. In scenario C, we assumed that the number of new partners was 0.5 for the peripheral group, and estimated the value of p_t to be 0.003. The formula shows that the consequent monthly HIV risk (p_i) is 0.000014. This value produces an accelerating endemic spread of HIV. Figure 6 illustrates that a combination of, for instance, n=5 and p_t=0.0003 gives the same value of p_i. In other words, the risk is the same if one simultaneously increases the safety in sexual practice and the number of new partners by a factor of ten.

Simulation further showed that if safety alone was increased by a factor of ten (p_t = 0.0003), the endemic spread of HIV would halt. In this case, p_i is 0.000001. The lower curve in Figure 6 illustrates this situation. The upper curve is the present homosexual situation. The lower curve shows the borderline for activity just sufficient to halt the spread of the disease. Instead of increasing the safety in sexual practice, one could reduce sexual activity by a factor of ten, and achieve the same result.

CONCLUSION

The upshot of all this is that if we assume the unlikely Scenario C, and the heterosexual population behaves 90% less risky than the homosexual population with regard to one of the risk factors, we will not have an endogenous spread among heterosexuals. More precisely, their combination of sexual activity and safety would have to be on the safe side of the lower curve in Figure 6. In the more likely Scenario D, it would be sufficient to behave just slightly less risky than the peripheral, male homosexual group in New York City. This general conclusion is supported by a Norwegian study using a more elaborate simulation model (Stigum *et al.* 1989).

ENDNOTES

1. Prison inmates are often moved to upstate New York prisons and may make the data less valid.

2. See Crawford *et al.* (1988) for a detailed model listing.

3. See Koopman *et al.* 1988 and Stigum *et al.* (1989) for a more thorough discussion.

4. This estimate is another example of how little we understand or know about sexual behavior and how this lack of information affects our attempts to model HIV/AIDS. The results of the Kinsey studies (1948) are relied upon by many mathematical and statistical modellers for generating the current size of the homosexual and bisexual population in the U.S. because current data do not exist. Unfortunately, there is little empirical evidence to support the extrapolations made from these studies completed more than four decades ago to our present environment.

REFERENCES

Bacchetti P. and A.R. Moss 1989. Incubation Period of AIDS in San Francisco. *Nature*; March.

Crawford, C.M. 1989. Public Policy in the Age of AIDS: Concerns About Sexual Behavior and Dating Activity. *Proceedings of the National HRMOB Conference*, November, Boston, MA.

Crawford C.M., G. DiFerdinando and F. Wenstop 1988. Dynamic Modelling of the Human Immunodeficiency virus Epidemic. Technical Report, Department of Human Service Studies, Cornell University, Ithaca, New York.

Darrow, W.W.; Echenberg, D.F.; Jaffe, H.W.; *et al.* 1987. Risk Factors for Human Immunodeficiency Virus (HIV) Infection in Homosexual Men. *American Journal of Public Health*, 77, 479-483.

Fauci, A.S. 1988. Receding Hopes of AIDS Vaccines. *Nature*, 333:699.

Fultz, P.N.; Mclure, H.M.; Swenson, R.B.; *et al.*
1986. Persistent Infection of Chimpanzees with Human T-lymphotropic Virus Type III: A Potential Model for Acquired Immunodeficiency Syndrome. *Journal of Virology*, 58:116-124.

Kaplan, H.B.; Johnson, R.J.; Bailey, C.A.; & Simon, W. 1987. The Sociological Study of AIDS: A Critical Review of the Literature and Suggested Research Agenda. *Journal of Health and Social Behavior*, 28, 140-157.

Kinsey, A.C.; Poneroy, W.D. and Martin, C.E., 1948. *Sexual Behavior in the Human Male*, Saunders: Philadelphia.

Koopman, J.; C. Simon, J. Jacquez, J. Joseph, L. Sattenspiel and T. Park, 1988. Sexual Partner Selectiveness on Homosexual HIV Transmission Dynamics. *Journal of Acquired Immune Deficiency Syndromes*; 1:486-504.

Martin, J.L. 1987. The Impact of AIDS on Gay Male Sexual Behavior Patterns in New York City. *American Journal of Public Health*, 77:578-581.

Stigum, H., J.K. Groennesby, P. Magnus, J.M. Sundet and L.S. Bakketeig 1989. The Potential For Spread of HIV in the Heterosexual Population in Norway, A Model Study. Department of Epidemiology, National Institute of Public Health, Oslo, Norway.

Winkelstein, W. Jr., Padian, Samuel M., *et al.* 1987. The San Francisco Men's Health Study: III. Reduction in Human Immunodeficiency Virus Transmission Among Homosexual/Bisexual Men, 1982-1986. *American Journal of Public Health*, 77: 685-689.

APPENDIX: New York City Data - AIDS Incidence

Year	Period	Old	New
1980	0	0	2
	1	0	1
	2	2	2
	3	0	1
	4	1	1
	5	2	3
	6	2	1
	7	0	1
	8	1	0
	9	2	1
	10	4	4
	11	2	2
1981	12	5	2
	13	5	5
	14	2	3
	15	10	7
	16	3	3
	17	7	6
	18	12	10
	19	14	9
	20	14	12
	21	12	13
	22	11	12
	23	14	11
1982	24	14	11
	25	16	18
	26	19	17
	27	26	21
	28	24	20
	29	20	13
	30	26	22
	31	31	27
	32	35	37
	33	32	28
	34	28	21
	35	30	26
1983	36	49	45
	37	36	38
	38	41	36
	39	54	50
	40	53	40
	41	68	59
	42	45	33
	43	40	30
	44	65	56
	45	48	49
	46	54	49
	47	73	65
1984	48	67	62
	49	82	72
	50	70	71
	51	78	71
	52	97	87
	53	95	80
	54	93	90
	55	112	95
	56	103	95
	57	99	89
	58	110	92
	59	97	72
	60	118	104
	61	124	112
	62	138	130
	63	120	99
	64	137	110
	65	127	120
	66	141	121
	67	149	129
	68	116	105
	69	136	122
	70	126	109
	71	127	125
1986	72	167	154
	73	164	136
	74	150	130
	75	143	134
	76	167	152
	77	180	173
	78	188	180
	79	172	176
	80	163	168
	81	173	178
	82	155	150
	83	153	175
1987	84	143	167
	85		167
	86		187
	87		161
	88		158
	89		163
	90		169
	91		171
	92		124
	93		171
	94		140
	95		173
	96		143
	97		175
	98		161
	99		157
	100		147
	101		187