Health Transition Review, Supplement 2 to Volume 7, 1997, 45–62

HIV prevalence and lifetime risk of dying of AIDS*

John Blacker and Basia Zaba

Centre for Population Studies, London School of Hygiene and Tropical Medicine

Abstract

The paper examines the relationship between the lifetime risk of dying from AIDS and HIV prevalence, using a female stable population model in which the epidemic has stabilized. In addition to prevalence, lifetime risk is determined by various other factors, notably the level of mortality from causes other than AIDS, age at infection, and survival time between infection and death. Typically, the lifetime risk of dying from AIDS is between three and five times the HIV prevalence. Regression equations are developed for estimating lifetime risk from the prevalence and other parameters. The methods are applied to data for Kenya, and it is shown that the 1995 prevalence estimate of 7.5 per cent for the population aged 15 and over would be equivalent to a lifetime risk of about 30 per cent.

HIV prevalence, or the proportion of the adult population which is HIV-positive, is the most commonly used index for measuring the scale of the epidemic. In African countries, national estimates of HIV prevalence are generally only available for females, as the main source of data is the anonymous testing of women who use antenatal services. More direct measures of the effect of HIV could be found from the annual death rate due to AIDS in the adult population, or the proportion of all adult deaths which are due to AIDS. These measures are rarely available in practice, as most African countries do not have death registration systems, and in any case, AIDS deaths are frequently perceived as due to other causes (Boerma et al. 1997). None of the above measures tells us about the overall risk of contracting HIV or dying from AIDS faced by an individual approaching adulthood and sexual maturity. But this last measure, which for a disease like AIDS is three or four times as large as the current prevalence measure, is one which would be very useful to planners and health educators alike. It would enable those who train skilled workers and professionals to judge what proportion of their cadres would be lost to this disease before the end of their working lives. It would allow insurance companies to set realistic premiums in countries affected by the epidemic. And it could provide a powerful advocacy tool for those trying to persuade young people not to follow the norm in terms of widely prevalent unsafe sexual behaviour.

HIV prevalence, or the proportion of the adult population which is HIV-positive, is the most commonly used index for measuring the scale of the epidemic. In African countries, national estimates of HIV prevalence are generally only available for females, as the main source of data is the anonymous testing of women who use antenatal services.

This paper investigates the relationship between HIV prevalence and lifetime risk, using models of age-specific incidence which have been tailored to reflect the type of sexual union formation patterns observed amongst females in a wide range of African populations. We focus on females because information on HIV prevalence, participation in sexual unions and fertility is widely available for females but not for males. The relationships which we discover

* Thanks for helpful comments on an earlier draft to Michael Bracher, Hans Oluf-Hansen, Martina Morris and Sam Preston.

in our model populations enable us to estimate lifetime risk statistics in populations which do not have the data required to calculate such risks directly.

Although HIV prevalence rates are usually quoted for 'the adult population' there is no generally accepted definition of 'adult', and different countries and different modellers adopt a variety of age ranges in their calculations. In Kenya, the National AIDS Control Programme defined prevalence relative to the population aged 15 and over, and the estimated figure for 1995 was 7.5 per cent, or about one person in 13. Confronted with this statistic, an uninformed reader might be tempted to infer that, if one person in 13 has the disease, the lifetime risk, or the proportion who could expect to contract it during the course of their adult lives, would also be about one in 13. In reality however the lifetime risk is very much higher than the prevalence, since the denominator for the prevalence rate consists of those currently alive, rather than all those who were ever at risk. On the other hand, detailed surveillance studies of small populations affected by HIV (e.g. Boerma et al. 1997; Nunn et al. 1997) have shown that at prevalence levels of between five and seven per cent, over 45 per cent of all adult deaths are due to AIDS. But it would be equally incorrect to assume that the lifetime risk of dying from AIDS for an adult in these populations was as high as 45 per cent, since the distribution of deaths by cause is affected by the age structure of the population, and in a growing population, younger adult ages, in which AIDS deaths predominate, would carry more weight.

Epidemiologists use a variety of measures to gauge the severity of a disease in a population. The most commonly used are prevalence: the proportion of the population suffering from the disease at a given point in time; and incidence risk: the annual number of new cases observed divided by the number of healthy individuals[1]. The relationship between prevalence and incidence is determined by the disease duration, which can be equated to survival time for diseases such as HIV/AIDS, which are generally fatal. In stationary and stable populations, simple relationships exist between these measures and the lifetime risk of dying of the disease (Preston 1987).

In a stationary population with a constant size and structure, and in which the number of births each year is exactly equal to the total number of deaths, these relationships are:

$$\frac{prevalence}{1-prevalence} = incidence \times mean\ survival\ time \qquad \ldots \mathbf{1}$$

$$lifetime\ risk = \frac{prevalence}{mean\ survival\ time \times birth\ rate} \qquad \ldots \mathbf{2}$$

For a stable epidemic in a stable population, with constant fertility, mortality, and growth rate, and a constant proportionate age structure, these equations are modified to give the following approximate relationships:

$$\frac{prevalence}{1-prevalence} \approx incidence \times mean\ survival\ time \times \exp(-rd) \quad \ldots \mathbf{3}$$

[1] Throughout this paper, all the incidence measures used are 'occurrences/exposure'-type measures, as is the convention in epidemiology, rather than 'occurrences/total population'-type measures which are occasionally referred to as demographic incidence rates, e.g. Hoem 1978.

$$lifetime\ risk \approx \frac{prevalence \times \exp(r \cdot [a+d])}{mean\ survival\ time \times birth\ rate} \qquad \ldots \mathbf{4}$$

where r is the growth rate of the population, a is the mean age at infection in the population, d is the mean length of time since infection for those people suffering from the disease, so that (a+d) is the mean age of persons currently infected[2]. When prevalence is defined as relating to a specified age group, then the 'birth rate' becomes the number of persons reaching the lower bound of the age group annually, divided by the total number in the age group. In populations which are growing in an irregular fashion, due to past changes in vital rates, the values *r* in equations 3 and 4 would be replaced by integral expressions of the type $r = \int_a^w r(x,t)dx$ appropriate for the time in question.

Data on all these variables are rarely available, but certain relationships are clear. Thus lifetime risk is higher when prevalence is higher; survival time is lower (this would also tend to shorten the mean time since infection); age at incidence is higher; mortality from other causes is lower (since growth rate will be higher). A low fertility rate might increase lifetime risk, as it would lower the crude birth rate in the denominator; but it would also reduce the growth rate in the numerator, so it could act in either direction.

Preston's equations allow us to appreciate the possible scale of the difference between prevalence and lifetime risk. Table 1 shows some hypothetical data conforming to the relationships shown in equation 4, for two different types of fatal disease. Both are assumed to have a prevalence of two per cent, a survival time of eight years, and a mean duration since infection of five years. But the cancer-type disease, which is shown occurring in a slow-growing population (growth rate 1%) has an older pattern of infection (mean age of incidence 50) than the AIDS-type disease (mean age of incidence 25) which is shown in a population growing at three per cent per annum. If it reached this prevalence level in spite of its old age pattern, the cancer-type disease would carry a lifetime risk of almost 20 per cent, nearly ten times its prevalence rate. A disease like AIDS which strikes at earlier ages, at this level of prevalence would be associated with a lifetime risk over five times its prevalence.

[2] Equations 3 and 4 would be exact if infection was concentrated at a single age, and if the illness was of fixed duration; a better approximation could be derived by using a series expansion of exp(r[a+d]) in terms of the means and variances of a and d.

Table 1
Theoretical relationship between disease prevalence and lifetime risk of dying (hypothetical data)

Disease	prevalence	Growth rate	Age at infection	Duration since infection	Survival time	Birth rate	Lifetime risk	Risk prevalence
Cancer	0.02	0.01	50.00	5.00	8.00	0.02	0.196	9.8
AIDS	0.02	0.03	25.00	5.00	8.00	0.04	0.114	5.7

An empirical example of a large difference between prevalence and lifetime risk is furnished by the example of maternal mortality. In the Matlab study area of Bangladesh, the maternal mortality ratio for the period 1976-89 was 5.1 maternal deaths per 1000 live births. The average total fertility rate during this period was 5.3, so the lifetime risk was about 2.7 per cent[3]. The incidence of maternal mortality may be taken as the maternal mortality rate of 91 per 100,000 woman-years (Fauveau and Chakraborty 1994). Assuming the average survival time for a fatal pregnancy to be nine months or 0.75 years, and the mean time since start of such a pregnancy for a cross-section of women to be half this figure, then with a growth rate of 1.3 per cent per annum, the prevalence of maternal mortality works out at 69 per 100,000 woman-years. Thus the lifetime risk of maternal mortality is nearly 40 times the prevalence of fatal pregnancies in this population.

In order to get realistic values for the age pattern of infection and subsequent mortality, we have used a spreadsheet model originally developed by Zaba (1994) for measuring the demographic effect of AIDS. This model allows for both behavioural and biological factors to influence the way HIV builds up in a cohort, and shows the effect of HIV on the growth and structure of the population. We define prevalence as the proportion HIV-positive among the population aged 15 and over. Lifetime risk is calculated by finding the number of deaths from AIDS in a typical cohort using multiple-decrement life table techniques, rather than the theoretical relationships above. The multiple-decrement approach eliminates the necessity of calculating statistics such as mean time since infection. A full formal specification of the model can be found in Zaba (1994), but a list of variable parameters which govern its operation is shown in Table 2. This table also shows the default values of the parameters and selected statistics for the default population and the epidemic so generated, as well as the ranges of the values used for the simulation models described later in this paper. In all our simulations we generate stable epidemics in stable populations, that is, the results presented show the relationships between prevalence and lifetime risk in a mature epidemic in which prevalence has reached an equilibrium level, and where death rates have stabilized along with the new population age structure.

[3] Estimated using the relationship:

$$lifetime\ risk = 1 - (1 - maternal\ mortality\ ratio)^{total\ fertility\ rate}$$

Table 2
Input parameters and output characteristics of simulation model

Union dynamics			
Input parameters	Default	Range	Notes
Women ever entering union	100%		
Sexual initiation:			
start age	15	13 - 17	Coale & McNeil (1972) marriage
median age	20	18 - 22	model
Union breakdown rate per year	10%	0 - 20%	Independent of age or union duration
Peak annual union re-entry rate	70%	50 - 90%	Bell-shaped distribution
Peak age for union re-entry	30	27 - 33	
Postpartum abstinence	6 months	3m - 9m	Range of variation: 0 - 2 years
Terminal abstinence median age	50	40 - 60	Logistic curve
Coital frequency			
new partner	0.3	.15 - .45	Average declines linearly from
regular partner	0.06	.03 - .09	once every 3 days for new unions to once every 2 weeks for unions lasting five years or more
Output characteristics			(pre HIV)
In active union by age 50	60%		
Mean partners by age 50	2.6		
Proportion partners dead	5%		determined by prevailing mortality rates
Fecundity			
Input parameters			
Mean age at			
menarche	13	11 - 15	Wood & Weinstein (1988) model
menopause	45	40 - 50	
Primary sterility	2%		
Maximum secondary sterility	40%	20 - 60%	Logistic curve
Breastfeeding amenorrhoea	6 months		Shortened after infant death
Postpartum contraception	1.5 yrs	0.0 - 3.0	

Table 2 (cont.)
Input parameters and output characteristics of simulation model

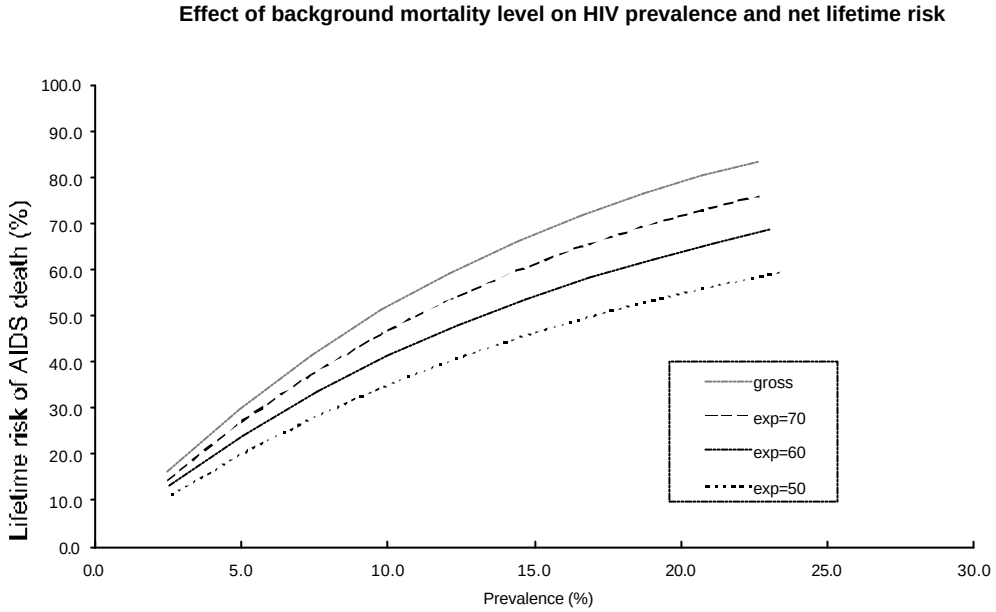
Output characteristics			(pre HIV)
Total fertility rate	5.00		jointly determined by fecundity and union dynamics parameters
Disease severity			
Input parameters			
Mean survival with HIV:			
adult	eight years	4 - 12	exponential decay curve
infant	1 year	0.5 - 1.5	all HIV+ve births die before age 10
Extent of vertical transmission	30%	20 - 40%	
Infection risk per coitus:			
new partner	0.0015	0 - .003	declines linearly over first five
regular partner	0.00015	0 - .0003	years of new partnership; range used in simulations 0.0 to 0.0003 for regular partners, new partner risk is ten times higher
Life expectancy of HIV-negative	60 years		Brass logit relational model
Output characteristics			(post HIV)
Stable prevalence at ages 15-50	17%		
HIV-positive births	6%		
Population life expectancy	39 years		
Mean age of infection schedule	32 years		

Mortality from causes other than AIDS

When discussing lifetime risks, it is necessary to distinguish between gross and net risks. A gross risk is the probability of a person contracting HIV, assuming that no one in the population dies from any other cause. A net risk is the proportion who will die of AIDS when allowance has been made for mortality from other causes. The difference between the gross and the net risks will of course depend on the level of mortality from causes other than AIDS; clearly the higher the mortality from other causes, the more likely will people be to die of something else first.

Since background mortality is a variable which can be manipulated in our model, we have calculated the net lifetime risks assuming that mortality from causes other than AIDS can be represented by model life tables with expectations of life at birth of 50, 60 and 70 years. The other variables were held constant at their default values shown in Table 2. The results are shown in Figure 1. The gross risk increases non-linearly with prevalence, and is over four times as high. The net risks are substantially smaller than the gross, but even with the highest level of background mortality, net lifetime risk is more than twice the prevalence. The lifetime risks shown in the next two figures are all net risks assuming an expectation of life of 60 years for causes other than AIDS.

Figure 1

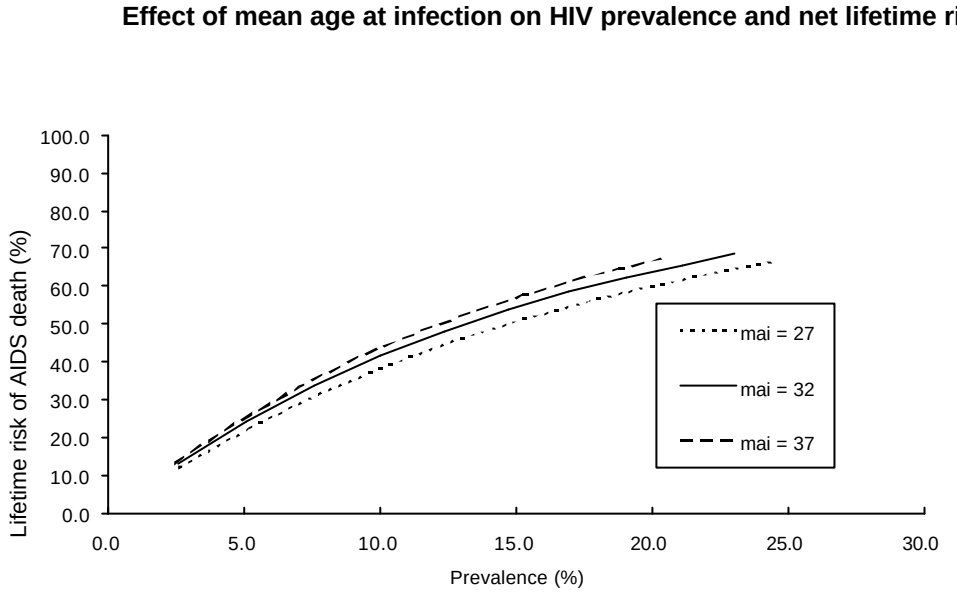


Age at infection

The mean age at infection is primarily determined by the ages at which people become sexually active. It thus varies from population to population, and between males and females, being generally younger among the latter. It is also affected by the age structure of the population, being younger when there is a steep slope on the age distribution. However this last effect can be eliminated from comparisons of populations if we make the calculations, not for the living population at a point in time, but for the schedule of incidence risks by age. The latter will normally be some four or five years older than the former.

Models with mean ages of the risk schedule of 27, 32 and 37 years were constructed. Survival time was held constant at eight years, and total fertility at five births per woman. The results are shown in Figure 2. The effects of the age at infection on lifetime risks are moderate but systematic. The higher the age at infection, the higher the lifetime risk implied by a given prevalence. This is because when age at infection is higher, the infection is concentrated in a smaller proportion of the population, so the same prevalence for the 15+ age range could only be achieved as the result of a higher overall lifetime risk.

Figure 2

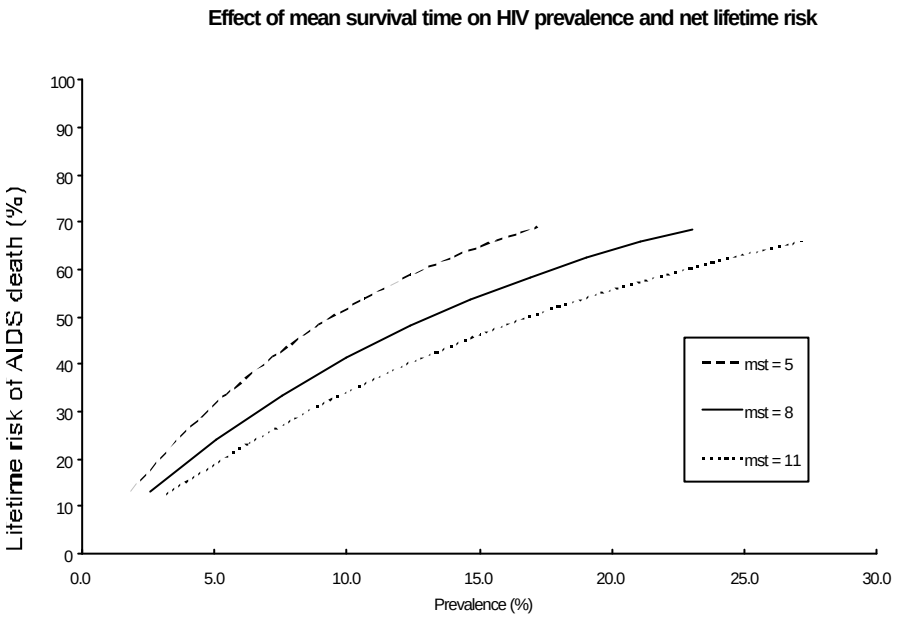


Survival time after infection

Mean survival times of the HIV-positive are difficult to estimate, since this involves very long prospective studies, but quite a lot of evidence is now available about median survival times. Cohort studies of white males in the US have suggested median survival times of ten years (Chin 1995), and drug therapies are now lengthening this time, but there is reason to suppose that in Africa it may be substantially less (Anderson 1989; Gregson 1994). Nunn et al. (1997) estimated median survival time after infection in Masaka district in Uganda as just over five years for persons infected when they were under age 55, and less than three years for those infected at later ages. In a simple exponential model of survival of the HIV-positive, such as used in our model, a mean survival time of eight years is equivalent to a median survival time of 5.5 years.

In order to explore the effect of survival time on the relationship between lifetime risk and prevalence, we have used three values of the mean duration: 5, 8 and 11 years. Mean age of the infection risk schedule was held constant at 32 years. The results are shown in Figure 3. Again the relationship is strongly affected by the survival time: for a given level of prevalence within our range, the lifetime risks were between 30 and 65 per cent higher when the mean duration was five years as against 11 years.

Figure 3

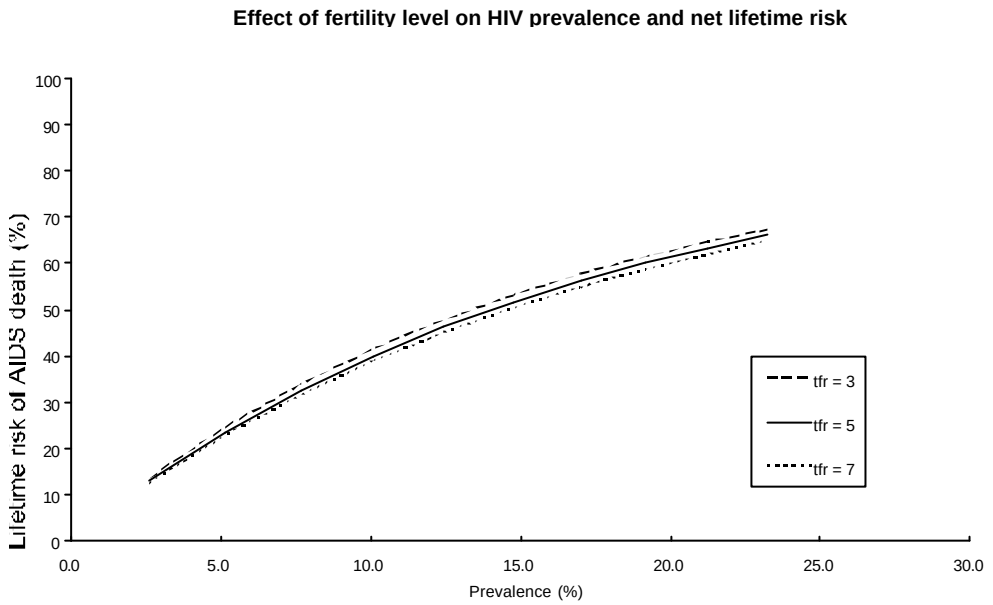


Fertility

As is evident from equation (4), fertility can affect the relationship between prevalence and lifetime risk in different ways. This is because fertility is the primary determinant of the shape of the age distribution: lowering fertility reduces the proportion of young people and increases the proportion of the old. With the mean age at infection held constant at 32 and the mean survival time at eight years, three sets of models were generated with total fertility rates of seven, five and three births per woman. The results are illustrated in Figure 4.

It appears that variation in the level of fertility, and hence in the age structure, has very little effect on the lifetime risk. This means that the relationship in populations with changing fertility levels will be similar to those observed in stable populations with an intermediate level of fertility. The explanation lies in the compensating effects of changes in the proportions of young adults under 25 and of old persons over 50; as one goes up the other goes down, but as both sections have relatively low prevalence of HIV the changes cancel out. If prevalence had been defined as the percentage HIV-positive among persons aged 15-49, a stronger relationship with fertility might have been revealed.

Figure 4



Child and adult risks

The lifetime risks shown in Tables 1 to 4 incorporate both the risks of horizontal transmission between adults and those of vertical transmission from mothers to children. The balance between adult and child risks depends on the likelihood of vertical transmission. In this model we have assumed a 30 per cent probability of an HIV-positive pregnant woman transmitting the virus to her unborn child, or during delivery or through breastfeeding, and we assume that all HIV-positive infants die before reaching adulthood. At this level, the risk of dying from HIV infection in childhood is between seven and ten per cent of the total risk, and the proportion increases slightly with the prevalence. If, however, we assumed a higher or lower vertical transmission ratio (VTR), the proportion of child deaths would increase or decrease correspondingly: with a VTR of 20 per cent childhood deaths constitute 4-6 per cent of total AIDS deaths; whereas if VTR was as high as 40 per cent, between 9 and 13 per cent of AIDS deaths would occur in childhood.

Thus the vertical transmission ratio is an important variable in determining the risk of infection in infancy and childhood, though not in adulthood. As such it is also a component of lifetime risk, albeit a relatively small one. It has virtually no effect on prevalence, which is based on the adult population only.

Applications to developing countries

The question now arises whether these models can be used to make realistic calculations of lifetime risk for Third World countries. Various problems arise in this respect.

In the first place information on at least two of the crucial variables, mean age at infection and survival time, is invariably lacking. We have attempted to overcome this problem by using proxy variables which we expect to be closely related to survival time and mean age at infection. We have investigated two such proxies which can be calculated from the models and which might also be available in real populations: the mean age at death of persons dying of AIDS, and, for females, the mean age at the birth of the first child. Sometimes information is available on the age distribution of AIDS cases rather than on deaths from AIDS, in which case the age at death can be approximated by adding half a year to the mean age of the AIDS cases, since the average survival time after the development of full-blown AIDS is only about one year; for example Boerma et al. (1997) report a mean duration of illness of ten months before an AIDS death.

In Africa AIDS cases and deaths from AIDS tend to be seriously underreported. But unless the degree of underreporting is associated with age, the mean age of the reported cases should give a reasonable approximation to that of all AIDS deaths. However bias could clearly arise if the degree of underreporting is greater in rural than in urban areas, and if the mean age at infection also differs between rural and urban areas.

The mean age of mothers at the birth of their first child should be associated with the mean age of HIV infection since both are the result of sexual activity. We would expect this relationship to be much closer in non-contracepting populations, where both pregnancy and infection are risks associated with sexual activity. But even in non-contracepting populations we would not necessarily expect the relationship to be fixed; if some cultures allow young girls to become sexually active before they become fully fecund, the risk of infection will be raised compared to that of pregnancy. And if sexual activity takes place with more than one partner, this will raise the risk of infection but not of pregnancy. Furthermore a high prevalence of primary infertility and high incidence of abortion would also raise infection risks relative to fertility risks in the population.

The mean age at first birth can readily be obtained from census or survey data on the distribution of women by parity and age group. Thus the mean age at first birth can be calculated from the proportions of childless women in each age group in the same way as the singulate mean age at marriage is obtained from the proportions never-married by age (Hajnal 1953). If there are appreciable numbers of women whose parity was not stated, adjustments by El Badry's (1961) method may be tried.

Information on median survival times may sometimes be available for African populations from small-scale longtitudinal studies, from which an informed guess can be made for the country concerned. On the assumption that the deaths are distributed in an exponential decay function, the mean values implied by the medians can be calculated. We have also explored the possibility of estimating survival times given both the mean age at first birth and mean age at death from AIDS, and have derived the necessary regression equations, using both ordinary least squares regression and a two-stage (instrumental variables) procedure, in which the estimate of mean survival time was not made explicit, but was indirectly fed into the regression estimate of lifetime risk. The results, however, were disappointing. The regression for mean survival based on the mean age at first birth, mean age at death from AIDS and prevalence explained only 67 per cent of its variance. The standard errors of the estimates of mean survival time were disturbingly large, giving 95 per cent confidence limits of the order of ±2 years when we are trying to estimate survival times of the order of eight years. When applied to the Kenya data, described below, they gave a figure of less than three years, which was unacceptably short. We have therefore been reluctantly forced to the conclusion that an independent estimate of survival time is an essential piece of information for our purposes, for which there is no readily available substitute.

A second problem with the practical applications is that the models are of stable populations with constant fertility, mortality and age structure, and in which the HIV/AIDS epidemic has also stabilized, so that prevalence too is constant. In practice we are dealing with situations where the epidemic is still growing and where both fertility and mortality from causes other than AIDS may be changing. In some countries badly affected by the AIDS epidemic, such as Zimbabwe and Kenya, fertility has been falling rapidly in recent years. Current growth rates will therefore differ from the stable population growth rates in our models, and even if recent growth rates can be estimated from successive censuses these will not persist, and would not be suitable inputs into the estimating procedure. These differences between the models and actual populations will affect our calculations in two ways: the underlying age structures will differ, and the mean age at infection will change as the disease spreads from age group to age group.

Changes in age structure will affect the prevalence, but not the lifetime risk, since the latter, like the expectation of life and the total fertility rate, is an index which is independent of the age structure. We have investigated the effect of a changing age structure on prevalence using the model fitted to the Kenya data described below. The age-specific prevalences from the female stable population model were combined with the projected female population of Kenya for 1995 (CBS 1996 b: appendix III); the overall prevalence for the population aged 15 and over worked out at 7.8 per cent as against 7.5 per cent in the model, which implied that we had overestimated lifetime risk by about one percentage point. Such an error can be regarded as trivial compared with the other margins of uncertainty surrounding our estimates.

If the mean age at infection rises as the epidemic spreads, the mean age at death from AIDS will also rise. But, because of the long period of incubation, the current mean age at death will imply an age at infection appropriate for a time several years before (which may partly explain why our regression estimates of survival time for Kenya were unacceptably short). Furthermore the mean age at death will also be affected by changes in the age structure of the population.

To assess the possible nature and magnitude of these changes, we have calculated, first, the mean age of the age-specific AIDS prevalence rates of the model projections made by Anderson (1989: Annex Table 14), and, second, the mean age of the AIDS cases obtained by multiplying their prevalence rates by their projected populations. Thus the first would measure the epidemiological changes only: the way the epidemic spreads from age group to age group; the second will measure the combined effects of the epidemiological and age structure changes.

The mean age of the AIDS prevalences rose by about two years over the 20-year period of the projections. In contrast the mean age of the AIDS cases rose by only 0.3 years over the same period. This implied that the epidemiological and age structure changes were having compensating effects. Their net result would be to introduce a small downward bias in the estimates of lifetime risk, thus compensating for the bias in the prevalences. We may conclude therefore that our stable population models can be used to estimate lifetime risk in current populations, provided they are not regarded as anything more than 'ball-park' estimates.

Linearization of the relationships and regression estimates

In order to make our model useful for estimating lifetime risk from commonly available data, we investigated the relationship between lifetime risk and our chosen indicators in a variety of circumstances. We generated 1000 model simulations with a wide range of values of HIV prevalence, non-AIDS mortality, mean age at infection, mean age at death from AIDS, mean age at first birth, survival time, vertical transmission ratio and fertility. The full model of HIV, population growth and structure is controlled by 22 different parameters describing biological and behavioural factors which govern fertility, mortality and HIV transmission. Even if each

parameter were only allocated one of two values, high and low, to produce simulations covering every possible combination of parameters would require 2^{22} = 4,194,304 different simulations. Since the amount of computer power needed to generate, store and analyse this would be excessive, an alternative strategy was adopted, whereby in each of the 1000 simulations each parameter was assigned a random value between its upper and lower limits.

The scatter diagram of the relationship between lifetime risk and prevalence in the female population derived from these models is shown in Figure 5. It will be seen that the trend is curvilinear, a feature which is already apparent in Figures 1 to 4. The curvature can however be eliminated if we plot the log-odds (i.e. the natural logarithms of the odds-ratios) of the lifetime risks against the corresponding log-odds of the prevalences, as shown in Figure 6. This linearization facilitates the use of multiple regression for analysing the effects of the different variables in the determination of lifetime risk. The independent variables, apart from prevalence, included mean age at first birth, total fertility, mean age at death from AIDS, mean survival time and the vertical transmission rate. An additional term of the product of the life expectancy of non-AIDS mortality and the log-odds of the prevalence was included, since the lines of the log-odds transformation of Figure 1, although straight, were not quite parallel.

The log-odds of the risks generated by the models were then regressed against the log-odds of the prevalences and the other variables. The regression analysis was carried out using ordinary least squares regression in the STATA (1997) package. Three separate regressions were made for each of the net risks considered as dependent variables: log odds of lifetime risk, log odds of risk for adults conditional on surviving to age ten, and the log odds of the net risk of dying from AIDS before age ten. The results are shown in Table 3. Apart from the independent variables listed in this table, we also experimented with the mean age of childbearing in the population, but in all cases this proved statistically insignificant, even at a ten per cent level, and was dropped from the final regressions. Unsurprisingly, the vertical transmission ratio proved to have an insignificant relationship with adult lifetime risk, and childhood risks were not significantly affected by the adult survival time after HIV infection, nor by the cross-product of life expectancy and prevalence. All the dependencies listed in Table 2 were highly significant at probability levels less than 0.1 per cent.

Figure 5

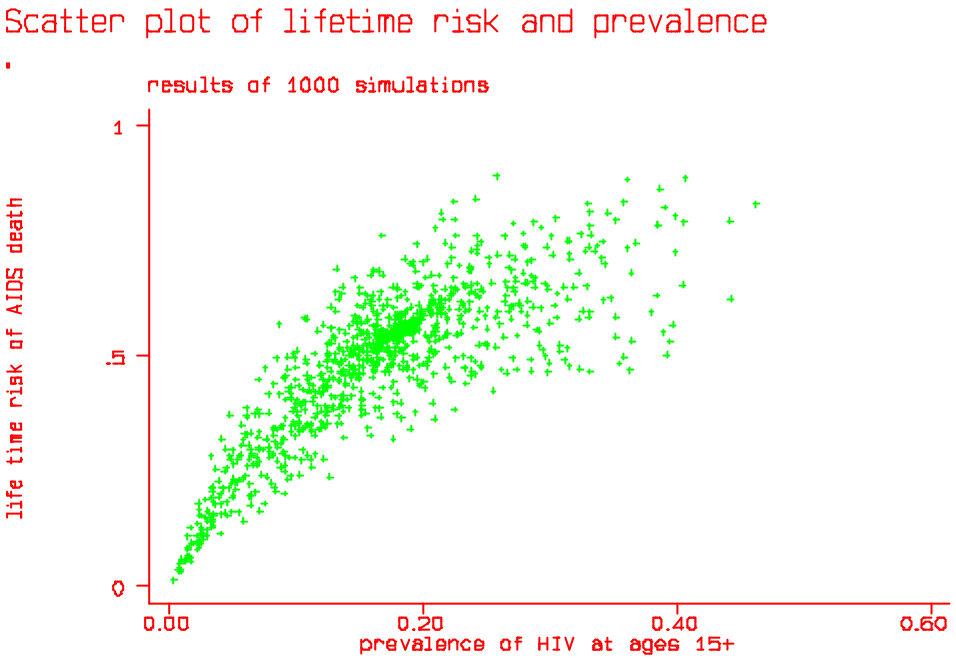
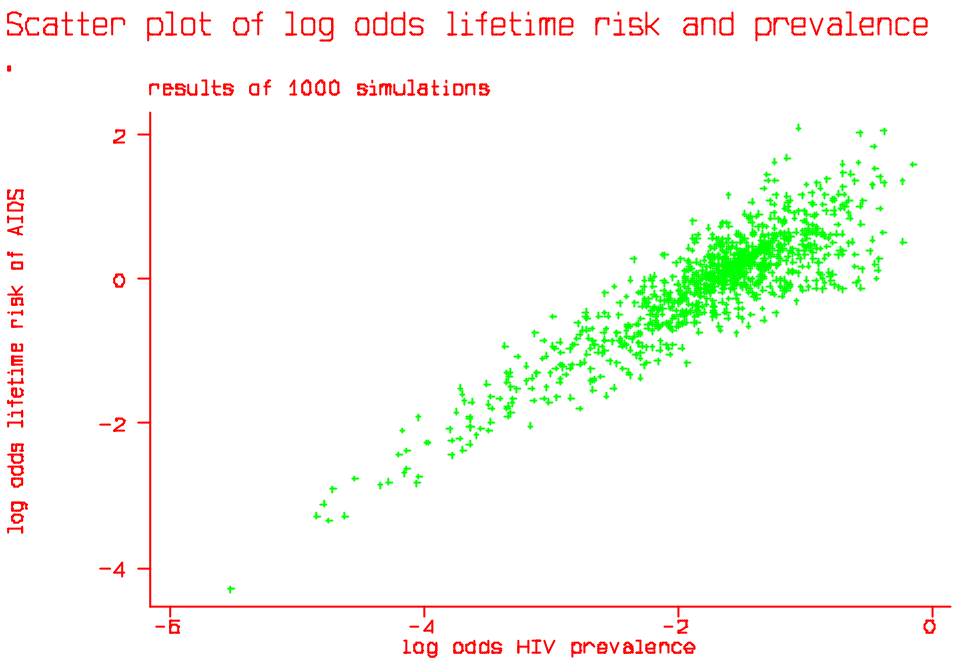


Table 3
Regression coefficients for the estimation of log-odds of net lifetime risk

	Risk at birth	Risk at age 10	Risk of dying from AIDS before age 10
Constant	-2.50	-1.43	-3.47
Log-odds prevalence	0.39	0.49	0.84
Life expectancy	0.06	0.05	0.01
Life exp.x log-odds prevalence	0.01	0.01	
Vertical transmission ratio	0.35		3.62
Mean age at first birth	0.03	0.03	0.07
Total fertility	0.03	0.02	0.06
Mean age at death from AIDS	0.03	0.03	-0.03
Survival time	-0.13	-0.15	
Regression statistics			
R-squared	0.99	0.99	0.99
Standard error	0.10	0.08	0.07

Figure 6

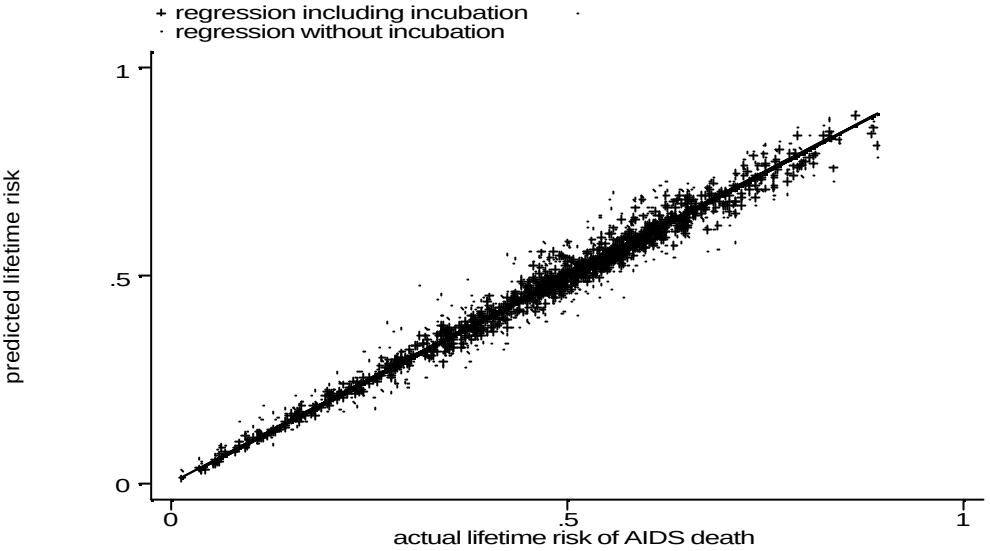


In the case of adult risk, and overall lifetime risk, the factors with the most influence were prevalence, its cross-product with life expectancy, life expectancy, survival time, and mean age at death from AIDS. Dropping the two factors relating to the age pattern of fertility would lower the amount of variance explained from around 99 per cent to 95 per cent, thereby raising the standard errors of the log odds estimates from around 0.09 to 0.17: this is equivalent to an error of 1-2 percentage points in the estimate of lifetime risk when this is of the order of 15 per cent.

Although survival time with HIV is one of the most important factors, both from theoretical considerations (equation 4) and as identified by the regression analysis above, we explored the effects of leaving it out of the regression procedure. The results showed that the amount of variance captured fell to around 90 per cent, and the standard errors of the log odds of the risks rose to 0.24, that is around four percentage points in estimating a risk of the order of 15 per cent. This we deemed unacceptable, which is why the results of regressions without survival time amongst the independent variables are not shown here. Figure 7 shows how well predicted values of lifetime risk based on regressions with and without survival time compare with the actual values generated in the model simulations.

Figure 7

Comparison of regression predictions and model output



A worked example: Kenya

As mentioned above, the National AIDS Control Programme of the Government of Kenya has estimated that the national HIV prevalence in 1995 was 7.5 per cent of the population aged 15 and over (NACP 1996). They have assumed a similar level of prevalence among males and females. They also estimate the vertical transmission ratio to be '30 to 40 per cent'; we have therefore adopted a value of 35 per cent. Figures of registered AIDS cases for the years 1988 to 1996 by sex and age group have also kindly been supplied by the NACP; from these we have estimated the mean age at death from AIDS for females to be 30.6 years.

The expectation of life for females in the absence of AIDS was taken as 61.4 years: this was the value obtained from the model life tables fitted to the data on child survival and orphanhood from the 1979 and 1989 censuses, and represented average values for the intercensal period, when mortality from AIDS could be regarded as negligible (CBS 1996a). Since mortality, both of children and adults, had been falling steadily before the 1989 census, it might seem reasonable to assume that mortality from causes other than AIDS would continue to fall into the 1990s. However this is a questionable assumption, and in the absence of hard evidence we have used the 1979-89 estimates as being the latest available. If, in so doing, we have overestimated non-AIDS mortality, we will have underestimated the lifetime risks of contracting HIV

The cohort mean age of women at the birth of their first child was calculated from the proportions of childless women in each age group (adjusted by the El Badry correction) obtained from the 1989 census, which gave a figure of 20.9 years. Similar calculations using the 1993 DHS data gave an identical result. Total fertility was taken to be 5.13, based on the the analysis of the 1989 census and 1993 DHS (CBS 1996b:13).

Our attempts to estimate mean survival time by regression gave unacceptable results. We know of no longitudinal studies in Kenya, but in Uganda and Tanzania such studies have suggested that seven years would be a plausible guess. The lack of information on survival times in Kenya constitutes the biggest gap in our knowledge for the purpose of these calculations.

When these parameters were multiplied by the regression coefficients shown in Table 3, this produced an estimate of overall lifetime risk of dying from AIDS of 31.3 per cent for a female born in Kenya subject to the current age-specific infection risks. Breaking down the risks separately for child and adult deaths, we find that the risk of dying from vertically transmitted AIDS is 3.6 per cent, and the risk of dying from AIDS for someone who survives to age ten is 32.8 per cent. The reason for the higher probability of dying from AIDS for those who survive to age ten than for the newborn is that infant and child mortality from other causes is also high in Kenya; in the absence of AIDS, we would expect about 12.4 per cent of female children to die before age ten in any case.

As an additional check on these regression results, we have constructed a single stable population model such that prevalence of HIV, non-AIDS mortality, mean age at first birth and at death from AIDS, total fertility rate, survival time and vertical transmission ratio were as close to the Kenya female values as possible. Table 4 shows the female stabilized age distribution, the age-specific prevalences, and the number of infected persons in each age group out of a total of 10,000; and the gross and net risks at different ages. The net lifetime risks at birth and at age ten from this best fitting model are about two percentage points lower than the regression-based estimates. But they confirm the general level of the lifetime risk estimate as around 30 per cent, some four times higher than the prevalence estimate.

Table 4
Age patterns of infection in the model for Kenya females

Age group	Age distribution	Prevalence (%)	Infected persons	Age	Net risk (%)	Gross risk (%)
0- 4	1565	1.50	23	0	29.20	39.40
5- 9	1338	0.00	0	5	29.90	36.80
10-14	1205	0.10	1	10	30.10	36.80
15-19	1077	4.40	47	15	30.30	36.20
20-24	912	12.80	117	20	29.00	33.00
25-29	741	14.40	107	25	24.10	27.00
30-34	604	11.70	71	30	18.10	20.20
35-39	501	8.70	44	35	12.80	14.10
40-44	422	6.70	28	40	8.50	9.30
45-49	357	4.70	17	45	5.10	5.50
50-54	305	2.50	8	50	2.60	2.80
55-59	261	1.20	3	55	1.20	1.40
60-64	220	0.60	1	60	0.60	0.70
65-69	179	0.30	1	65	0.30	0.30
70-74	138	0.10		70	0.10	0.20
75+	175	0.10	0	75	0.00	0.10
Total	10000	4.70	468			
Total 15+	5892	7.50	443			

Table 4 explains this feature which was noted at the beginning of this paper: why the lifetime risk is so much higher than the prevalence. The latter is calculated over the whole 15+

age range, while the infections are concentrated in a relatively short age span. Thus prevalence is low at the extremes of the distribution: in the 15-19 age group which contains many who have not yet been infected but may well be later, and at higher ages (45 and over) where the survivors escaped infection while many of their contemporaries perished and have therefore disappeared from the prevalence calculations.

We may conclude therefore that the estimated prevalence for 1995 of 7.5 per cent, or one person in 13, implies that, of the children born in Kenya, almost one in three will die of AIDS, leaving the other two to die of other causes. Although our calculations cannot do more than indicate a rough order of magnitude, it is nevertheless a formidable figure which we feel deserves wider recognition.

References

Anderson, R.M. 1989. Mathematical and statistical studies of the epidemiology of HIV. *AIDS* 3, 6:333-346.

Boerma, T.J., J. Ngalula, R. Isingo, M. Urassa, K.P. Senkoro, R. Gabone and E.N. Mkumbo. 1997. Levels and causes of adult mortality in rural Tanzania with special reference to HIV/AIDS. This volume.

Central Bureau of Statistics (CBS). 1996a. *Kenya Population Census 1989: Analytical Report.* Vol.3, *Population Dynamics of Kenya*. Nairobi.

Central Bureau of Statistics (CBS). 1996b. *Kenya Population Census 1989: Analytical Report,* Vol. 7, *Population Projections.* Nairobi.

Chin, J. 1995. Scenarios for the AIDS epidemic in Asia. Asia- Pacific Population Research Reports No.2. Honolulu: East-West Center.

El Badry, M.A. 1961. Failure of enumerators to make entries of zero: errors in recording childless cases in population censuses. *Journal of the American Statistical Association* 56,296:909-924.

Fauveau, V. and J. Chakraborty. 1994. Women's health and maternity care in Matlab. Pp. 109-138 in *Matlab: Women, Children and Health*, ed. V. Fauveau. Dhaka: International Centre for Diarrhoeal Disease Research, Bangladesh.

Gregson, S. 1994. Will HIV become a major determinant of fertility in sub-Saharan Africa? *Journal of Development Studies* 30,3:650-679.

Hajnal, J.1953. Age at marriage and proportions marrying. *Population Studies* 7,2 111-136.

Hoem, J. 1978. Demographic incidence rates. *Theoretical Population Biology* 14:329-337.

National AIDS Control Programme (NACP). 1996. *AIDS in Kenya*. Nairobi: Ministry of Health.

Nunn, A.J., D.W. Mulder, A. Kamali, A. Ruberantwari, J.Kengeya-Kayondo and J. Whitworth. 1997. Five-year HIV-1 associated mortality in a rural Ugandan population. *British Medical Journal*, in press

Preston, S.H. 1987. Relations among standard epidemiologic measures in a population. *American Journal of Epidemiology* 126,2:336-345.

STATA. 1997. Statistical Software, Release 5. Stata Corporation.

Zaba, B. 1994. The demographic impact of AIDS: some stable population simulation results. Research Paper 94-2. London: Centre for Population Studies, London School of Hygiene and Tropical Medicine.