An introduction to the mathematics of HIV/AIDS modelling

Leigh Johnson
Centre for Actuarial Research
June 2004

The purpose of this note is to introduce the key concepts in the mathematical modelling of HIV/AIDS. The note starts with a brief introduction to epidemiological and modelling terminology. It then presents a basic model of HIV transmission, and discusses how this model can be improved to allow for greater accuracy in modelling of HIV transmission and survival under HIV infection. Finally, the estimation of numbers of orphans is discussed.

1. Epidemiological terminology

Epidemiology is the study of the distribution and determinants of diseases – both infectious and non-infectious diseases. Originally the term was used to refer only to the study of epidemic infectious diseases, but it is now applied more broadly to other diseases as well. Questions such as, 'How much does smoking increase the risk of lung cancer?' and 'How much does cholesterol increase the risk of heart disease?' would therefore be included within the ambit of modern epidemiology. Epidemiologists also concern themselves with questions of disease control: to what extent do preventative interventions reduce the incidence of disease, and to what extent do therapeutic interventions lead to cure or symptom relief?

In examining the distribution of disease, it is important to distinguish between prevalence and incidence. The **prevalence** of a disease is defined as the percentage of a particular population that is infected with the disease. The **incidence** of the disease is the rate at which new infections occur. Hence, if the number of people infected with a particular disease at the start of a particular year is d, out of a population of size N, and n new infections occur over the course of the year, then the prevalence rate at the start of the years is d/N and the annual incidence rate is $n/(N-d)$. When a disease is introduced into a population, it usually expands rapidly at first, with prevalence rates rising. However, as people either die from the disease or develop resistance to it, the pool of susceptible individuals shrinks, and the prevalence starts to decline. In some cases, the prevalence drops to zero, but in many cases, prevalence stabilizes at a non-zero level that is referred to as the **endemic prevalence**.

2. Modelling terminology and types of models

A key conceptual distinction is that between models based on population averages and individual-based simulations. In **individual-based simulations**, each individual in the population is modelled as a discrete entity, and characteristics are determined separately for each individual. In models based on **population averages**, it is assumed that all individuals in the population have identical characteristics. The most common modelling approach is in fact somewhere between individual-based simulations and population averages; most models divide the population into **cohorts** of individuals, and individuals in the same cohort are assumed to all have the same characteristics. The characteristics used to define cohorts in models of HIV/AIDS are usually factors such as age, sex and level of sexual risk behaviour. Cohorts are also usually defined according to disease status: susceptible, infected (possibly further split according to stage of disease) or resistant.

Stochastic models are often used with individual-based simulations. These models allow events such as HIV infection and death to be simulated by random processes. **Deterministic models**, however, calculate *expected* numbers of events in cohorts of individuals. They are therefore used with models based on population averages, or models that are based on the division of the population into cohorts of individuals with the same characteristics. Deterministic models generate unique solutions, because they are based only on *average* values of random processes. A stochastic model, however, generates different solutions each time it is run, because the answers depend on the actual simulation of the random processes.

Suppose, for example, that there are 10 individuals with AIDS, and the probability of death is 0.4 per annum. In a deterministic model, the number of AIDS deaths in the next year would be calculated as 4 (0.4×10). In a stochastic model, however, the number of deaths would be assumed to follow a binomial distribution, with parameters 10 and 0.4. To simulate the number of deaths, we take the inverse of the cumulative distribution function of the binomial distribution and apply it to a randomly generated number between 0 and 1. If this random number were 0.15, for example, the simulated number of deaths would be 2. The random number changes each time the model is run, so that different results are obtained each time the model is run. The key advantage of a stochastic model is that if it is run a large number of times, it can give the user a sense of the range of uncertainty around the model outputs. A deterministic model can only tell the user what outcome is expected on average.

We can also distinguish between compartmental and distributional models. In **distributional models**, individual characteristics are treated as continuous variables that can take on any value within a specified range. In **compartmental models**, individual characteristics are treated as discrete variables, which can only take on a finite number of values. For example, in a distributional model, an HIV-positive individual's CD4 count might be allowed to take on any value between 0 and 2000. In a compartmental model, however, individuals would be classified according to the range in which their CD4 count falls: CD4 > 500, CD4 < 200, or CD4 between 200 and 500, for argument's sake. Models

that are based on population averages or division of the population into cohorts are by definition compartmental models, while individual-based simulations can be distributional or compartmental.

Events such as infection and death can be modelled as occurring over a time interval or at a point in time. In **discrete time** models, numbers of events are calculated over each time period. The time period can be a year, a month, a day or any other chosen frequency. The choice of time period will depend in practice on the type of phenomenon that is being modelled. If, for example, we are modelling the risk of HIV transmission during acute HIV infection, when viral load changes very rapidly over the course of a few weeks, it will be necessary to use weekly or daily projection intervals. If, however, we were modelling HIV survival in adults, it would be acceptable to use yearly projection intervals, because the time from HIV infection to death is usually around 10 years, and the mortality rate does not change substantially from one year to the next. In **continuous time** models, events are modelled as occurring at a point in time. These models are specified in terms of differential equations. The discrete time model is often an approximation to the continuous time model, and may require simplifying assumptions (for example, in deriving multiple decrement tables, actuaries often make the assumption that events are uniformly likely to occur over the course of a given year). Generally, the shorter the time interval in the discrete model, the closer the model results will be to those of the continuous time model.

Often, **analytical solutions** can be obtained to continuous time models. This means that the results of the model can be expressed algebraically, in terms of equations involving the key parameters. For example, in the next section, we show how the endemic prevalence of a disease can be calculated in terms of the model parameters if we are using a very simple mathematical model. When continuous time models are complex, however, it is often necessary to use **numerical methods**, such as the Runge-Kutta method (Press *et al,* 1986), which approximate the exact solution. With discrete time models, it is usually only possible to obtain **numerical solutions** i.e. the results of the model have to be obtained through an iterative calculation process.

Some modellers also draw a distinction between curve-fitting models and simulation models. In a **curve-fitting model**, the user enters the HIV prevalence rates and the model determines what HIV incidence rates are most consistent with these prevalence rates. An example of such a model is the Epidemic Projection Package (EPP) model, used by UNAIDS. In this model, HIV incidence is modelled by a function with four parameters, and a least squares procedure is used to determine what combination of parameters gives the greatest level of consistency with the HIV prevalence data. In a **simulation model**, assumptions are entered about sexual behaviour patterns and HIV transmission probabilities, and these determine the HIV incidence rates.

For further explanation of these and other terms, see Garnett (2002).

3. A basic model of an HIV/AIDS epidemic

Consider a sexually active population that is divided into two groups: susceptible (i.e. HIV-negative) individuals and infected individuals. Let the number of susceptible individuals be X and the number of infected individuals Y. Let N be the sum of X and Y, i.e. the total sexually active population. It is assumed that π is the number of individuals entering the sexually active population per period, and individuals leave the sexually active population at a rate μ (in the absence of AIDS). It is assumed that β is the probability of HIV infection in a single sexual partnership with an HIV-positive individual, and c is the rate at which new partnerships are formed. It is further assumed that once people become infected with HIV, they experience an increase in the rate of mortality, equal to α. These dynamics are illustrated in Figure 1 below.

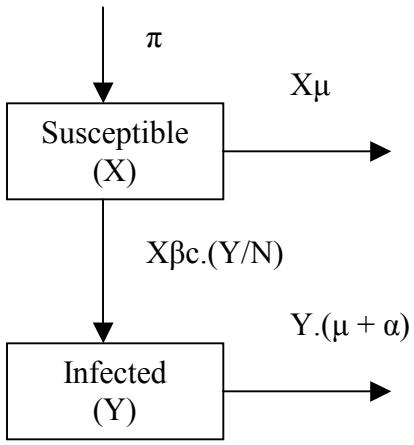

Figure 1: A basic model of HIV transmission and survival

The rate at which new infections occur is calculated as the rate at which new partnerships occur (c), multiplied by the probability of infection if the partner is HIV-positive (β), multiplied by the probability that the partner is HIV-positive (Y/N). The model can also be expressed in terms of differential equations, as follows:

$$\begin{aligned} \frac{dX}{dt} &= \pi - X\beta c.(Y/N) - X\mu \\ \frac{dY}{dt} &= X\beta c.(Y/N) - Y(\mu + \alpha) \end{aligned} \qquad (1)$$

This is probably the most basic model of HIV transmission and survival that can be constructed. In terms of the terminology explained in the previous section, the model can be described as

- based on population averages (all individuals are assumed to have the same sexual behaviour patterns);
- deterministic (no allowance is made for random variations);
- set in continuous time (differential equations are being used);
- having analytical solutions; and
- based on simulation rather than curve fitting.

A concept that is often used in this type of basic epidemiological model is the concept of the **basic reproductive number**, or $\boldsymbol{R_0}$. The basic reproductive number is defined as "the average number of secondary cases of infection generated by one primary case in a susceptible population" (Anderson *et al,* 1995). In other words, if there was only one infected person in a large population, R_0 is a measure of how many other people that person would infect before they either died or ceased to be infectious. R_0 is the product of the probability of transmission per unit time and the duration of the period of infectiousness. The probability that an infected individual transmits the virus in a given unit of time (Δt) is $c.\beta.\Delta$t, and the average duration of infectiousness, in the units of time measured, is $1/\left((\mu+\alpha)\Delta t\right)$. R_0 can therefore be calculated as

$$c.\beta.\Delta t \times \frac{1}{(\mu+\alpha)\Delta t} = \frac{c.\beta}{\mu+\alpha}. \qquad (2)$$

R_0 is a measure of how transmissible a disease is; the higher it is, the more rapidly the disease spreads and the higher its endemic prevalence. If R_0 is less than 1, this implies that for every new infection, less than one subsequent infection occurs. This implies that the disease will eventually die out. Much of the disease modelling literature has therefore focused on the question of how effective interventions such as vaccination and antiretroviral treatment need to be in order to reduce the basic reproductive numbers of different diseases to less than 1.

The basic reproductive number is also important because it can be used to calculate the endemic prevalence of the disease of interest. Suppose the size of the sexually active population is constant i.e. there are always exactly the same number of people entering the population as there are leaving. In this situation, a constant prevalence can only be maintained if every new infection generates one further infection, on average. The number of new infections generated by one infection is calculated as R_0, reduced in proportion to the number of sexual contacts who are susceptible to infection i.e. R_0 (X/N). For the equilibrium condition to hold, it must be the case that

$$\begin{aligned} 1 &= R_0\left(X/N\right) \\ \frac{1}{R_0} &= \left(1-\frac{Y}{N}\right) \\ \frac{Y}{N} &= 1-\frac{1}{R_0} = 1-\frac{\mu+\alpha}{c.\beta} \end{aligned} \qquad (3)$$

(Y/N) is the endemic prevalence. Clearly, the higher the basic reproductive number, the higher the endemic prevalence of the disease is likely to be. It is thus possible to obtain an analytical solution for the endemic prevalence in this basic epidemiological model.

This basic epidemiological model has the advantage of being simple and easy to understand, but it does also have several limitations. Firstly, the model does not allow for heterogeneity in sexual behaviour; it assumes that all individuals change partners at the same rate and that all infected individuals are equally likely to transmit the virus. The basic reproductive number becomes more difficult to calculate when heterogeneity in susceptibility and infectiousness are allowed for. Another problem is that the model is not structured by age, and it is therefore of limited use in long-term population projections. Many of the parameters in the model should, in theory, be allowed to vary with age; non-AIDS mortality, AIDS mortality, and rate of partnership formation are all examples. Thirdly, the model assumes a constant rate of AIDS mortality, α. Evidence suggests that the rate of mortality experienced by HIV-positive individuals increases as the duration of infection increases (UNAIDS Reference Group on Estimates Modelling and Projections 2002).

Lastly, some modellers prefer to use a rate of transmission per act of intercourse than to use a rate of transmission per partnership, because the risk of transmission depends to some extent on the amount of sex that occurs within a partnership. It should be emphasized, however, that much of the risk of HIV transmission is concentrated in the first few sexual contacts with the infected partner, and the risk of transmission per sex act is therefore not constant with respect to the cumulative number of sexual exposures to the infected partner (Downs and De Vincenzi 1996). Some modellers therefore argue that using probabilities of transmission per sex act is not necessarily more accurate than using probabilities of transmission per partnership (Garnett and Anderson 1993).

4. Improving the modelling of HIV transmission

In this section, possible improvements to the model of HIV transmission presented in section 3 are discussed. These changes are presented for a model that is set in discrete time, using annual projection intervals. This is done for the sake of consistency with the ASSA2002 model, on which much of this discussion is based.

The first improvement considered is the replacement of the rates of transmission per partnership with rates of transmission per sexual contact. Suppose that individuals have m sexual partnerships per annum, and that the average number of acts of intercourse per partnership is n. Further assume that the probability of transmission in a single act of sex is r, and the probability that a randomly selected partner is HIV-positive is p. Then the probability of becoming HIV-infected in a given year is

1 – Probability of not getting infected in the year
= 1 – (Probability of not getting infected by a randomly infected partner)^m
= 1 – (Probability that randomly selected partner is HIV-positive ×

Probability of not getting infected by an HIV-positive partner +
Probability that randomly selected partner is HIV-negative)^m

$$= 1 - \left(p(1-r)^n + (1-p)\right)^m \qquad (4)$$

Equation (4) can be further modified to allow for heterogeneity in sexual behaviour. Suppose the sexually active population is divided into four sexual activity classes, or 'risk groups'. The probability that an individual in risk group *i* becomes infected in a single year is then

$$1 - \left(\sum_{j=1}^{4} w_{ij} . p_j (1 - r_{ij})^{n_{ij}} + \left(1 - \sum_{j=1}^{4} w_{ij} . p_j \right) \right)^{m_i} \qquad (5)$$

where

w_{ij} is the proportion of the individual's partners who are in risk group *j* if the individual is in risk group *i*

p_j is the probability that a partner in risk group *j* is HIV-positive

r_{ij} is the probability that an individual in risk group *i* becomes infected through a single act of sex with an HIV-positive partner in risk group *j*

n_{ij} is the average number of sexual contacts per partnership in partnerships between individuals in risk groups *i* and *j*

m_i is the number of partnerships per annum for an individual in risk group *i*

Equation (5) is very similar in form to equation (4). The first important difference is that the *m*, *n*, *p* and *r* parameters in equation (4) have been changed to be dependent on the risk group of the individual and/or the risk group of the partner. The second important difference is that the summations across the possible partner risk groups have been inserted, with each term in the summation weighted by the w_{ij} term. Effectively, we are still trying to calculate the probability that a randomly selected partner does not infect the individual, but we are now calculating this probability separately for each possible partner risk group, and weighting this by the proportion of partners in each risk group.

The w_{ij} terms define the patterns of sexual mixing between the risk groups. A number of possible approaches can be taken in defining these terms. One possible approach is to define the values of w_{ij} separately for each possible combination of *i* and *j*, subject to the constraint that

$$\sum_{j=1}^{4} w_{ij} = 1 \qquad \text{for all values of } i \qquad (6)$$

This is the approach that is currently followed in the ASSA2002 model. It has the advantage of allowing great flexibility in terms of the modelling of sexual mixing

patterns. However, because the w_{ij} terms are fixed, this approach effectively assumes that sexual mixing patterns do not change over the course of an epidemic. It is likely that sexual mixing patterns would change over the course of an epidemic, as the relative sizes of the different risk groups change. Risk groups in which rates of partner change are high are likely to diminish dramatically as a result of AIDS mortality, so that relatively fewer partners are selected from these groups as the epidemic matures.

A more common approach to defining the sexual mixing patterns involves the setting of a single parameter, ε. The w_{ij} terms are then defined by equation (7) below (Garnett and Anderson 1996):

$$w_{ij} = (1-\varepsilon)\delta_{ij} + \varepsilon\left(\frac{N_j.m_j}{\sum_u N_u.m_u}\right) \qquad (7)$$

where

$\delta_{ij} = 1$ when $i = j$ and 0 when $i \neq j$

N_j is the number of individuals in risk group j

m_j is the average annual number of partnerships for an individual in risk group j

ε is a parameter which can take on any value between 0 and 1. When ε is 0, w_{ij} is equal to 1 when $i = j$, and 0 otherwise. This means that individuals are assumed only to have sex with other people in the same risk group. This special case, in which individuals have strong preferences about the risk group of their partners, is referred to as **completely assortative mixing**. At the other extreme, when ε is 1, the first term on the right-hand side of equation (7) falls away, and the probability that the partner is in risk group j depends on the number of individuals in risk group j and the average number of partners they have. In this situation, individuals do not have any preferences about the risk group of their partner, and the w_{ij} terms become independent of i. This is referred to as **random mixing**.

Some sense of the significance of the ε parameter can be gained from Figure 2. This shows the endemic prevalence of a sexually transmitted disease (STD), as a function of a measure of STD transmissibility, for different mixing parameters. (The 'mixing parameter' referred to in this diagram is equal to $1 - \varepsilon$.) The figure demonstrates that the closer the mixing patterns are to completely random, the higher the endemic prevalence of the STD is likely to be. This occurs because as people become less selective in choosing their partners, it becomes easier for the STD to spread beyond the high-risk groups.

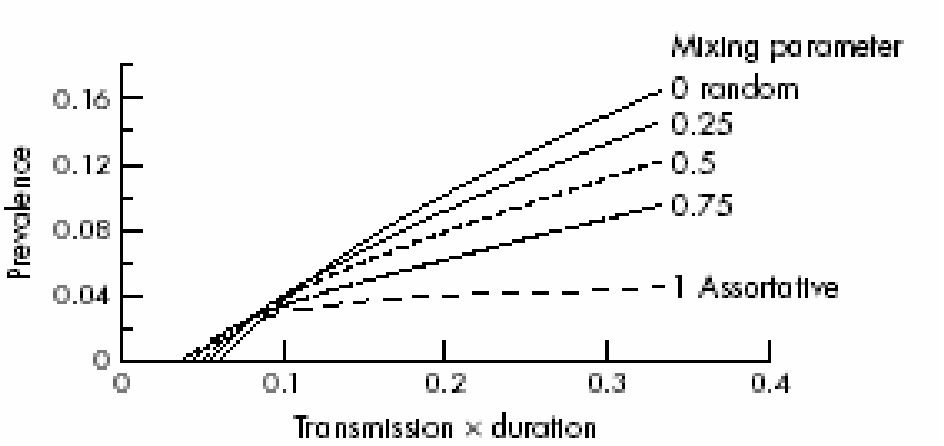

Figure 2: Relationship between endemic prevalence and degree of sexual mixing
Source: Garnett (2002)

The model of HIV transmission can be further improved to take into account the effect of age on the risk of infection. The probability that an individual of age x, in risk group i, becomes infected in a given year is

$$1-\left\{\sum_{j=1}^{4} w_{ij} \sum_{y=14}^{59} h(y \mid x).p_j(y).(1-r_{ij})^{n_{ij}s(y)} + \left(1-\sum_{j=1}^{4} w_{ij} \sum_{y=14}^{59} h(y \mid x).p_j(y)\right)\right\}^{m_j s(x)} \qquad (8)$$

where

$h(y \mid x)$ is the proportion of partners who are aged y if the individual is aged x

$p_j(y)$ is the probability that a partner in risk group j, of age y, is HIV-positive

$s(y)$ is an adjustment to allow for the effect of age on the frequency of sexual behaviour

Equation (8) is similar to equation (5). Again, summations across the possible partner ages have been added, with each term in the summation weighted by the $h(y \mid x)$ term. The probability that the individual is not infected by a partner is now calculated separately for each possible combination of partner age and partner risk group, and weighted according to the proportions of partners of the relevant age and risk group.

A number of further improvements could be made to this model. Thus far, we have been considering a single-sex population. This may be appropriate if one were modelling a mainly homosexual epidemic, e.g. the early AIDS epidemic in San Francisco. In modelling a heterosexual epidemic, however, it is helpful to divide the population into males and females. The male-to-female probability of HIV transmission per sexual contact is higher than the female-to-male transmission probability, and HIV prevalence rates therefore tend to be significantly higher in women than in men in heterosexual epidemics. In addition, high-risk behaviour tends to be more concentrated in women than in men, with a small group of women (mainly sex workers) having extremely high rates of partner change, while the rest of the female population have low rates of partner

change. Dynamics such as these play an important role in the evolution of an AIDS epidemic.

The model of HIV transmission can also be extended to allow for the reduction in transmission risk due to condom use, and the increase in transmission risk due to other STDs. Some models make an implicit allowance for the effects of STDs by increasing the HIV transmission probabilities in those risk groups in which STD prevalence is likely to be highest (e.g. Dorrington (2000)). Other models simulate the transmission and cure of other STDs explicitly, so that the increase in HIV transmission risk can be modelled more accurately (e.g. Korenromp *et al* (2002)). However, most models of the interaction between HIV and other STDs examine the effects of only one or two other STDs, and therefore do not capture the full complexity of the relationship between HIV and other STDs.

Lastly, the model of HIV transmission can be further extended to allow for the changing risk of HIV transmission over the HIV-positive partner's course of disease. The viral load changes significantly between initial infection and death, and has been shown to be highly predictive of the risk of HIV transmission (Quinn *et al,* 2000). The viral load reaches high levels during the first few weeks of infection, then drops to low levels, and then gradually increases over the subsequent course of infection. As with most other STDs, though, the increase in infectiousness per sex act tends to be associated with an increase in the severity of the symptoms. As symptoms become more severe, individuals are likely to engage in sex less frequently, and this to some extent offsets the increased risk of transmission per sex act. It is therefore not always clear that symptomatic individuals are more infectious than asymptomatic individuals (Garnett and Bowden 2000).

5. Improving the modelling of HIV survival

One of the key concepts underlying the modelling of HIV survival is the concept of a **survivor function**. The survivor function, $S(t)$, can be defined as the number of individuals surviving to time t, where t, in our context, is the number of years since HIV infection. A number of different distributional forms for the survivor function can be used. Some of the most commonly used are listed below.

Weibull assumption (parameters λ and ϕ)

$$S(t) = \exp(-\lambda t^{\phi}) \qquad (9)$$

Exponential assumption (parameter λ)

$$S(t) = \exp(-\lambda t) \qquad (10)$$

Gompertz assumption (parameters B and c)

$$S(t) = \exp(-B(c^{t} - 1)/\ln c) \qquad (11)$$

Gamma assumption (parameters α and λ)

$$S(t) = 1 - \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \int_0^t x^{\alpha-1} e^{-\lambda x} dx \qquad (12)$$

The **Weibull** and **exponential** distributions are the distributions most commonly used in AIDS modelling. From equations (9) and (10), it should be clear that the exponential distribution is in fact the same as the Weibull distribution when ϕ is set to 1. More detail is presented on the Weibull distribution below, and the points made apply equally to the exponential distribution.

An alternative parameterization of the Weibull distribution, which is more commonly used, is given in equation (13) below.

$$S(t) = 0.5^{(t/m)^{\phi}} \qquad (13)$$

The **median**, m, is the time to which half the individuals survive i.e. of those individuals alive at time 0, half survive for less than m years, and half survive for more than m years. The UNAIDS Reference Group for Estimates, Modelling and Projections (2002) currently recommend a median of 9 years in modelling the time from HIV infection to AIDS-related death. It can be shown that the median, expressed in terms of the parameters in equation (9), is

$$m = \left(\frac{1}{\lambda} \ln 2 \right)^{1/\phi} \qquad (14)$$

Substituting equation (14) into equation (13) gives the same parametric form as in equation (9).

The **shape parameter**, ϕ, determines the variance of the time to death i.e. it affects how concentrated the distribution of survival times is around the average term to death. A commonly used shape parameter in the modelling of HIV survival is 2.5 (Gregson *et al,* 1998). Figure 3 below assists in the interpretation of the shape parameter. Three Weibull survivor functions are shown, each with a median of 10 years, and with different shape parameters. Because the median is the same for all three functions, they all pass through the same point, corresponding to the 50% survival rate. Mortality rates are constant when the shape parameter is set to 1 (the exponential assumption), but if the shape parameter is greater than 1, mortality rates increase as the duration of infection increases.

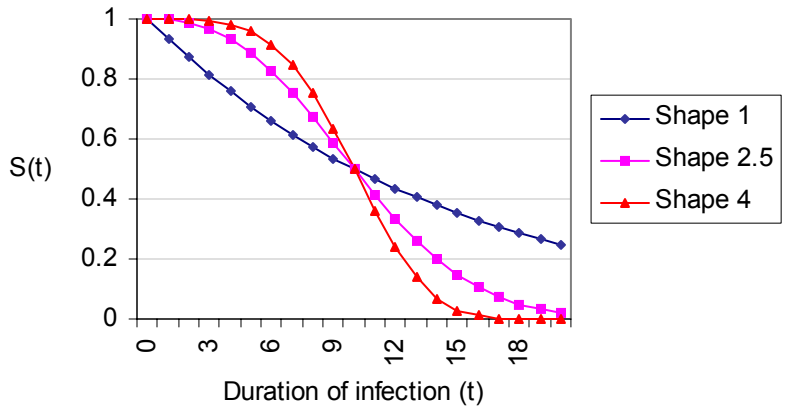

Figure 3: Weibull survivor functions with shape parameters of 1, 2.5 and 4

When modelling survival in children, it is necessary to allow for the fact that children infected at or before birth progress to AIDS and death much more rapidly than children infected after birth through breastfeeding. The approach to modelling paediatric survival that is recommended by the UNAIDS Reference Group on Estimates, Modelling and Projections (2002) is to use a 'double Weibull' distribution i.e. a survivor function of the form

$$S(t) = 1 - \theta\left(1 - \exp(-\lambda_1 t^{\phi_1})\right) - \left(1 - \theta\right)\left(1 - \exp(-\lambda_2 t^{\phi_2})\right) \qquad (15)$$

This survivor function is the weighted average of two Weibull survivor functions – one for children infected at or before birth, the other for children infected through breastfeeding. θ determines the weight given to the survivor function for children infected at or before birth, and hence can be thought of as the proportion of all children infected by their mothers who are infected at or before birth. t here is the child's age rather than the child's duration of infection, since most of the breastmilk transmission occurs during the first year of life, and in practice it is difficult to determine the exact timing of HIV transmission in children. The 'double Weibull' function currently used by UNAIDS in modelling paediatric survival is shown in Figure 4 below. The initial rapid decline in the proportion surviving is due to the high AIDS mortality experienced by children infected at or before birth. With fewer of these children remaining in the subsequent years, mortality rates start to fall, then rise again at older ages as more of the children infected through breastmilk start to die.

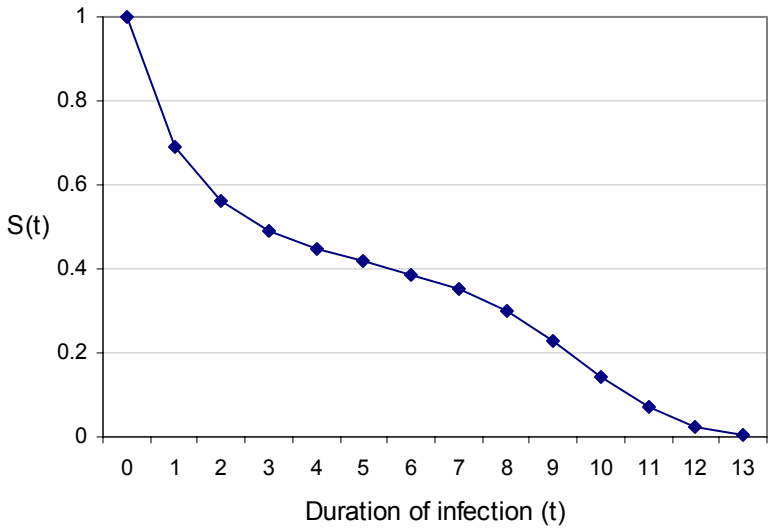

Figure 4: Paediatric HIV survival, using a 'double Weibull' survivor function

6. Estimating numbers of maternal orphans

This section and the next section discuss possible approaches to estimating numbers of orphans. Various definitions of the term 'orphan' are possible. In this section, we consider the estimation of **maternal orphans** i.e. children who have lost a mother or a mother and a father. **Paternal orphans** are similarly defined as children who have lost a father or a mother and a father. **Dual** or **double orphans** are children who have lost both parents. The estimation of dual and paternal orphan numbers is more difficult than the estimation of maternal orphan numbers, and is discussed in the next section.

The method described here for estimating orphan numbers is based loosely on work by Timæus and Grassly (2001). Figure 5 may assist in the interpretation of the symbols defined below.

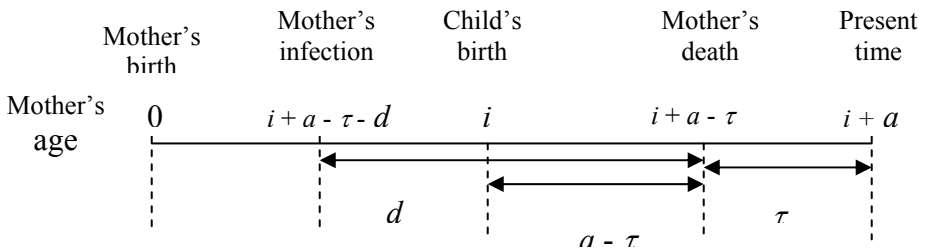

Figure 5: Time line representing age and time variables

Note that it is possible for the mother to be infected after the child's birth (so that $d < (a - \tau)$), or for the mother not to be infected at all.

We define the following symbols:

$\Omega_t(a,\tau)$ = number of maternal orphans of age *a* at time *t* whose mothers died τ years ago

Ω_t = number of maternal orphans at time *t*

$$\Omega_t = \sum_{a=0}^{14} \sum_{\tau=0}^{a} \Omega_t(a,\tau) \qquad (16)$$

(This assumes we are defining maternal orphans to be under the age of 15 – the upper limit on the first summation can be changed to reflect other age definitions).

$\mu'_{t-\tau}(d,r,i+a-\tau)$ = deaths among women in risk group *r*, who were HIV+, with duration *d* years since infection at death, aged ($i + a - \tau$), who died in year ($t - \tau$)

$\mu_{t-\tau}(r,i+a-\tau)$ = deaths among women in risk group *r*, who were uninfected at the time of death, aged ($i + a - \tau$), who died in year ($t - \tau$)

$m'_{t-a}(i,d,r)$ = fertility of woman in risk group *r*, aged *i*, duration *d* since infection, at time ($t - a$)

$m_{t-a}(i,r)$ = fertility of woman in risk group *r*, aged *i*, uninfected, at time ($t - a$)

ξ_p = probability of transmission from HIV positive mother to her child at/before birth

ξ_m = probability of HIV positive mother infecting her child through breast-milk

$p_{t-a}(a)$ = probability of uninfected child born at time ($t - a$) surviving to age *a*

$p'_{t-a}(a)$ = probability of child infected at/before birth, born at time ($t - a$), surviving to age *a*

$p''_{t-a}(a)$ = probability of child infected by breast milk, born at time ($t - a$), surviving to age *a*

It follows that the number of maternal orphans of age *a* at time *t*, whose mothers died uninfected, τ years ago, is

$$\sum_{r=1}^{4} \sum_{i=15}^{49} \mu_{t-\tau}(r,i+a-\tau).m_{t-a}(i,r).p_{t-a}(a) \qquad (17)$$

The number of maternal orphans of age *a* at time *t*, whose mothers died HIV positive, τ years ago, but became positive after giving birth, is

$$\sum_{r=1}^{4} \sum_{i=15}^{49} \sum_{d=0}^{a-\tau-1} \mu'_{t-\tau}(d,r,i+a-\tau).m_{t-a}(i,r).p_{t-a}(a) \qquad (18)$$

The number of maternal orphans of age *a* at time *t*, whose mothers died HIV positive, τ years ago, and became positive before giving birth, is

$$\sum_{r=1}^{4} \sum_{i=15}^{49} \sum_{d=a-\tau}^{29} \mu'_{t-\tau}(d,r,i+a-\tau).m'_{t-a}(i,d-a+\tau,r).$$

$$\left[\ \xi_p . p'_{t-a}(a) + (1-\xi_p)\xi_m . p''_{t-a}(a) + (1-\xi_p)(1-\xi_p)p_{t-a}(a)\ \right] \qquad (19)$$

It then follows that

$$\Omega_t(a,\tau) = \text{sum of (17), (18) and (19)} \qquad (20)$$

By substitution of equation (20) into equation (16), the total number of maternal orphans at time *t* can then be estimated. Equation (17) can be taken as an estimate of the number of 'non-AIDS orphans', while the sum of equations (18) and (19) can be taken as an estimate of the number of 'AIDS orphans'.

7. Estimating numbers of paternal and dual orphans

Three different approaches have been developed for the purpose of estimating paternal and dual orphan numbers.

The first method, currently used by UNAIDS, was developed by Timæus and Grassly (2001). Paternal orphan numbers are estimated using the same method as used for maternal orphans, with male fertility and mortality rates being used in place of female fertility and mortality rates. The important complicating factor is that allowance needs to be made for the possible discordance between the maternal and paternal HIV status i.e. if the father is HIV-positive, it does not necessarily imply that the mother is HIV-positive, and similarly, if the father is HIV-negative, it does not necessarily imply that the mother is HIV-negative. The percentage of children who are dually orphaned is then estimated as

% of children maternally orphaned × % of children paternally orphaned + excess

The 'excess' term in this equation represents the extent to which maternal and paternal mortality are correlated. If it is 0, maternal and paternal mortality are independent of one another. Timæus and Grassly estimate the 'excess' parameter by conducting a regression on Demographic and Health Survey (DHS) data from various African countries, in which percentages of children orphaned maternally, paternally and dually are estimated empirically. Factors such as HIV prevalence and proportions of married women in monogamous unions are found to estimate the 'excess' parameter, and are used to determine separate 'excess' risks of orphanhood for each country. This approach is described in more detail by the UNAIDS Reference Group on Estimates, Modelling and Projections (2002).

A second possible approach also involves the use of DHS data to estimate paternal and double orphan numbers. The method used by Hunter and Williamson (2000), in calculating the orphan estimates of the US Census Bureau, again assumes that estimates of maternal orphans are already available. It further assumes that the ratios of maternal: double orphans and maternal: paternal orphans are the same across all countries with generalized AIDS epidemics, but changing over time, as the AIDS epidemic matures.

These ratios can be estimated from DHS data. Once these ratios are estimated, the numbers of paternal and double orphans can be calculated from the numbers of maternal orphans, divided by the relevant ratios.

The third approach does not require the use of DHS data. This method, described by Johnson and Dorrington (2001), calculates the number of dual orphans by taking the number of new maternal orphans in each year, from women in each risk group, and projecting forward paternal mortality. Similarly, paternal orphans are estimated by taking the number of births in each year, from women in each risk group, and projecting forward paternal mortality. Male partner mortality, for women in risk group i, is calculated as the weighted average male mortality rate in each male risk group, where the weights are the w_{ij} terms defined in section 4, and the male mortality rates in each risk group are calculated by projecting the model forward. The advantage of this approach is that it is not reliant on DHS data, which could potentially be biased. Its disadvantage, however, is that it is largely dependent on the assumptions made about sexual mixing patterns, which are largely unknown. In addition, the method is computationally very complex.

References

Anderson R., Swinton J. and Garnett G. (1995) Potential impact of low efficacy HIV-1 vaccines in populations with high rates of infection. *Proceedings of the Royal Society of London, Series B.* 261: 147-151

Dorrington R. E. (2000) *The ASSA2000 suite of models*. Actuarial Society of South Africa Convention, 2000. Somerset West, South Africa. Available: www.assa.org.za/default.asp?id=1000000086. Accessed 25 February 2004

Downs A. and De Vincenzi I. (1996) Probability of heterosexual transmission of HIV: relationship to the number of unprotected sexual contacts. *Journal of Acquired Immune Deficiency Syndromes.* 11: 388-395

Garnett G. (2002) An introduction to mathematical models in sexually transmitted disease epidemiology. *Sexually Transmitted Infections.* 78: 7-12

Garnett G. and Anderson R. (1993) Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. *Philosophical Transactions of the Royal Society of London. Series B.* 342: 137-159

Garnett G. and Anderson R. (1996) Sexually transmitted diseases and sexual behaviour: insights from mathematical models. *Journal of Infectious Diseases.* 174(Suppl 2): S150-S160

Garnett G. P. and Bowden F. J. (2000) Epidemiology and control of curable sexually transmitted diseases: opportunities and problems. *Sexually Transmitted Diseases.* 27(10): 588-599

Gregson S., Machekano R., Donnelly C., *et al.* (1998) Estimating HIV incidence from age-specific prevalence data: comparison with concurrent cohort estimates in a study of male factory workers, Harare, Zimbabwe. *AIDS.* 12: 2049-2058

Hunter S. and Williamson J. (2000) *Children on the Brink: Executive Summary, Updated Estimates and Recommendations for Interventions*. United States Agency for International Development. Available: http://www.usaid.gov/pubs/hiv_aids/childrenreport.pdf

Johnson L. and Dorrington R. (2001) *The impact of AIDS on orphanhood in South Africa: a quantitative analysis*. Centre for Actuarial Research. Available: www.commerce.uct.ac.za/care/Monographs/Monographs/mono04.pdf

Korenromp E. L., Bakker R., Gray R., *et al.* (2002) The effect of HIV, behavioural change, and STD syndromic management on STD epidemiology in sub-Saharan Africa: simulations of Uganda. *Sexually Transmitted Infections.* 78(Suppl i): 55-63

Press W. H., Flannery B. P., Teukolsky S. A. and Vetterling W. T. (1986) *Numerical Recipes (Chapter 15: Integration of ordinary differential equations)*. Cambridge, Cambridge University Press

Quinn T., Wawer M., Sewankambo N., *et al.* (2000) Viral load and heterosexual transmission of human immunodeficiency virus type 1. *New England Journal of Medicine.* 342(921-929)

Timæus I. and Grassly N. (2001) *Orphanhood in populations with generalised AIDS epidemics: methods and results*. International Union for the Scientific Study of Population XXIV General Population Conference. Salvador, Brazil.

UNAIDS Reference Group on Estimates Modelling and Projections (2002) Improved methods and assumptions for estimation of the HIV/AIDS epidemic and its impact: recommendations of the UNAIDS Reference Group on Estimates, Modelling and Projections. *AIDS.* 16(WHO-UNAIDS Report): W1-W14