

5 Application to Uganda: Inputs

To test our “minimalist” model we have constructed a projection for Uganda. Uganda was chosen because it has one of the longest running severe HIV epidemics, and reasonably good data (though some of it based on small, localized studies) on prevalence trends, age patterns of incidence and survival after infection. In this chapter we describe the input data – sources, assumptions and adjustments – and discuss some of the results obtained with the projection model. Input schedules are illustrated graphically in the text, and the single year, numerical data are given in full in the Excel workbook *UgandaIn.xls*, which is available on the project web site. This should facilitate different implementations of our model, and will allow the Uganda data to be used as a substitute for local data in populations which are not so rich in information. Some of the computations which produce the numerical inputs for the projection are also available as separate workbooks on the web site, but not all of these are fully documented.

Most of our data come from published sources, or from publicly accessible data-bases such as those maintained by the UN Population Division, DHS (Macro International, <http://www.macrint.com/dhs/>) or the US Bureau of the census (<http://www.census.gov:80/ipc/www/hivaidn.html>). However, some of the information presented here from longitudinal studies of HIV affected populations has been made available directly to the UNAIDS reference group on estimates and projections prior to publication.

The 50 year projection period used in this worked example represents approximately 20 years historical experience of the epidemic, and 30 years extrapolation into the future. National level prevalence data are available for each calendar year from 1995 to 1998, based on anonymous unlinked screening of women attending ante-natal clinics.

Mortality of the uninfected

A Brass general standard lifetable was used to represent age specific mortality from non-HIV related causes, which was assumed constant for the projection period. The level parameter for females was set to -0.15 , for males it was set to -0.05 , giving pre-HIV life expectancies of 48.6 and 45.4 for females and males respectively, which are the values for life expectancy in Uganda for 1975-80 according to the UN World Population Prospects, 1998 revision (UN 1999). The force of mortality in a single year is estimated from the logarithm of the ratio of adjacent survivorship values. The single-year Brass standard lifetable and the calculation of the of the usual lifetable functions and the force of mortality based on these are shown in the *mu* worksheet of *UgandaIn.xls*. Figures 1 and 2 show plots of the survivorship functions and the forces of mortality by sex. The survivorship functions for males and females are distinct, but the forces of mortality for the sexes are undistinguishable on this scale.

Figure 1

Uganda, 1975-80 - survivorship pre-HIV

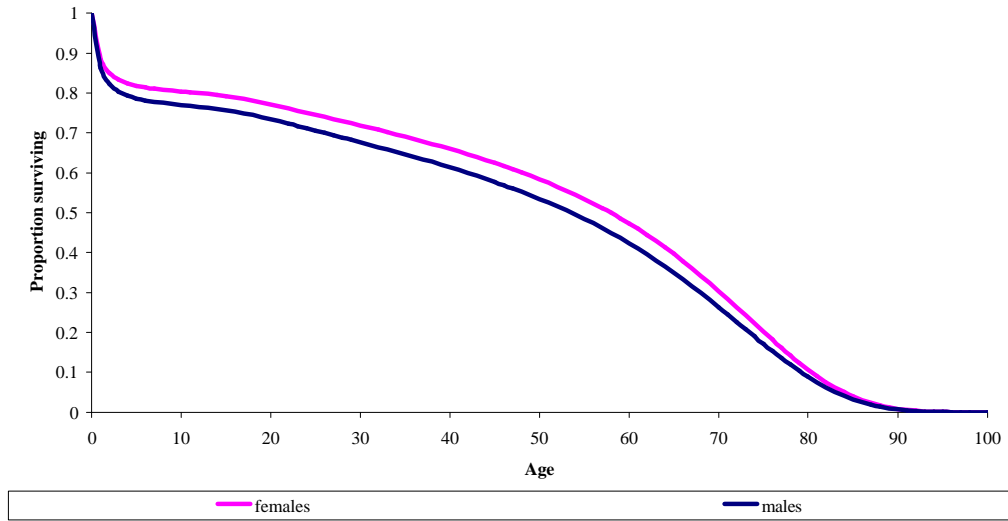
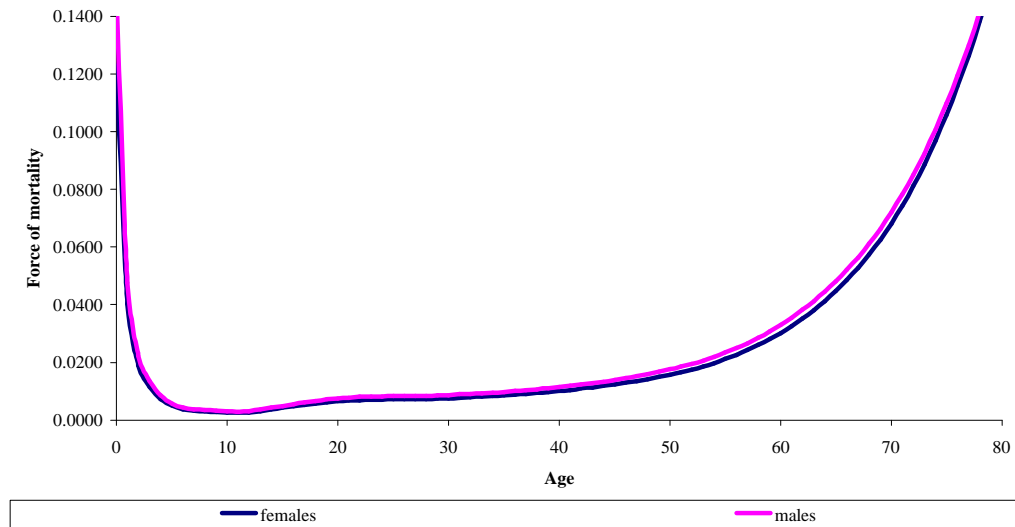


Figure 2

Uganda, 1975-80: force of mortality for non-HIV causes



Mortality due to HIV related causes

A preliminary analysis of survival data for persons infected with HIV by duration since infection and by age has been made available for this project from the Masaka MRC study site in Uganda. These data are based on follow-up of the natural history cohort – a subgroup of their general population cohort for whom age at infection is known to within a year, as sero-conversion occurred between two survey rounds at which testing took place. Tabulations were made for broad age groups, and for the two sexes separately. Cox regression showed a highly significant effect of age on survival ($p=0.0003$). Using the under 25 age group as baseline - the middle age group has 1.6 the hazard of death and the oldest have 8.7 the hazard (Whitworth, 1999, personal communication). However no independent effect of sex was found. We have therefore used this mortality data for both sexes specific for age and duration since infection in constructing smoothed lifetable survivorship functions to represent an “HIV only” mortality schedule from which we derived the forces of infection.

The Masaka data are shown in table 1. Figure 3 compares the observed data with model values generated using a suitably scaled cosine function to represent survival in the “AIDS only” lifetable. Numerical values of the single year age and duration specific force of mortality due to HIV are shown in the *alpha* worksheet of *UgandaIn.xls*. A separate worksheet, *mortalph.xls*, as yet only partially documented, is available showing how the model values are computed.

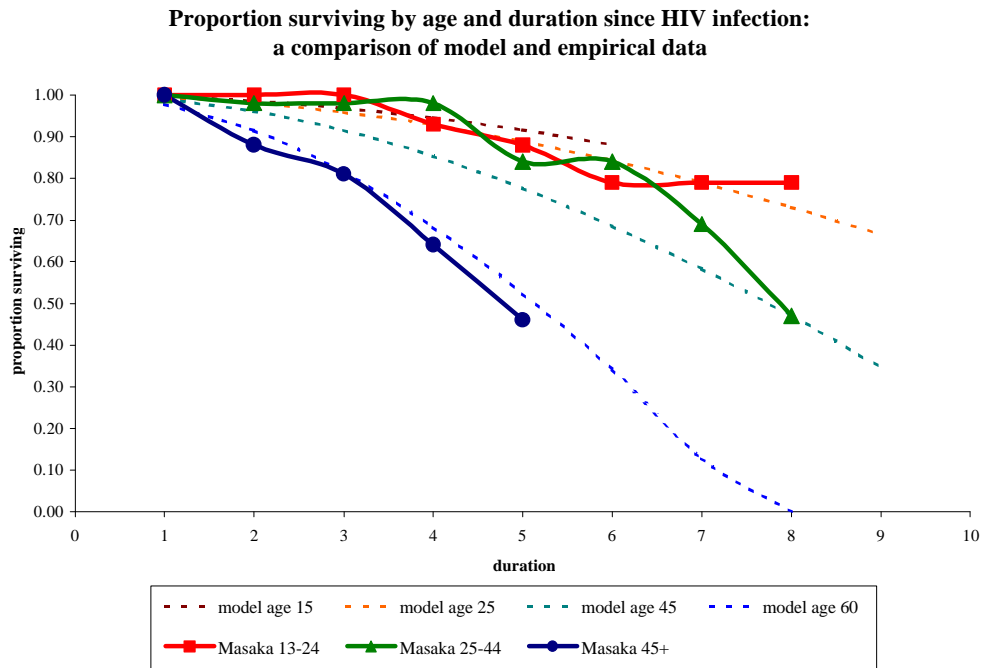
Table 1 *Percent surviving by duration since infection and current age, Masaka, 1990-98, both sexes*

Years since infection	Current age			
	under 25	25-44	over 45	all ages
1	100	100	100	100
2	100	98	88	98
3	100	98	81	97
4	93	98	64	92
5	88	84	46	81
6	79	84		74
7	79	69		67
8	79	47		57

need to get PYO basis for these figures from Jimmy – my guess is less than 400 PYO in total

The Masaka survival data are based on all deaths in the follow-up period, no attempt has been made in table 1 to separate deaths from HIV related causes from mortality from other causes. However, with fairly small numbers in this follow-up group (around 350 PYO at all ages), and with a crude adult death rate of around 8 per thousand we would expect fewer than 30 such deaths in the 8 year follow up period. There would be little point in attempting to assign the expected deaths to duration by age categories and performing a complex competing risk analysis. We note from figure 2 that the observed survivorship trajectories generally lie along the lower bound of the “HIV only” survivorship curves for the corresponding age groups in the model life tables. Adding mortality from other causes to the model would lower the model curves in relation to the observed data and thereby place the observed curves in a more central position within the appropriate model range.

Figure 3



Sexual activity forces

The Uganda DHS collected data on sexual activity, and from these data we can calculate the proportions ever sexually active by age, and these can be used, in theory, to estimate the age specific forces for becoming sexually active, and for ceasing sexual activity. More information is available for women than for men, since more women (7,000) were interviewed than men (2,000), women answered more detailed questions and more tabulations are provided for females in the DHS report. In estimating rates of entry into sexual activity, we were able to re-analyse individual level data on age at first sex (by single years of age) for women, but for men we only had current status data (ever had sex) in five year age groups from published tables.

To estimate the force of becoming sexually active for women, we first need to construct a life table showing proportions who have never had sex. There are two possible approaches – one would be to simply use the current status data, computing proportions at each age who report that they have never had sex. The other approach is to use retrospective reports of age at first sex, allowing for censoring of exposure to risk time in the younger members of the population. If we lump together the experience of women of all ages in either method, this is tantamount to ignoring any secular trend in age at sexual debut, though the second method should allow us to investigate this to a limited extent. Only the retrospective data on age at first sex allow us to estimate the magnitude of the force at ages under 15, as the DHS only covered women aged 15 to 49. As a check on the quality of the data we have used retrospective reports on age at first sex to generate our life table, and then compared this with current status data from age 15 onwards.

Figure 4

Lifetable for never sexually active - Uganda DHS data

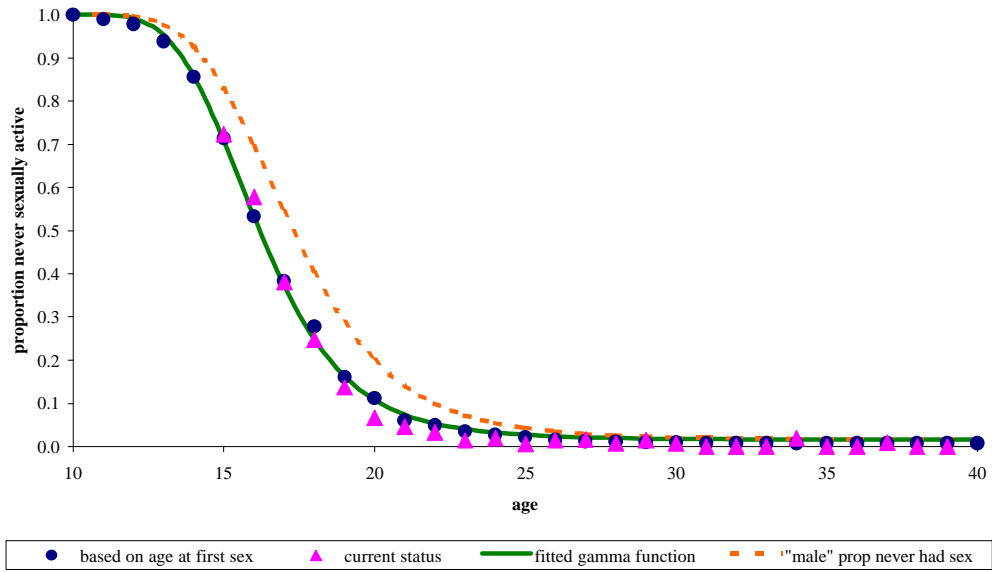
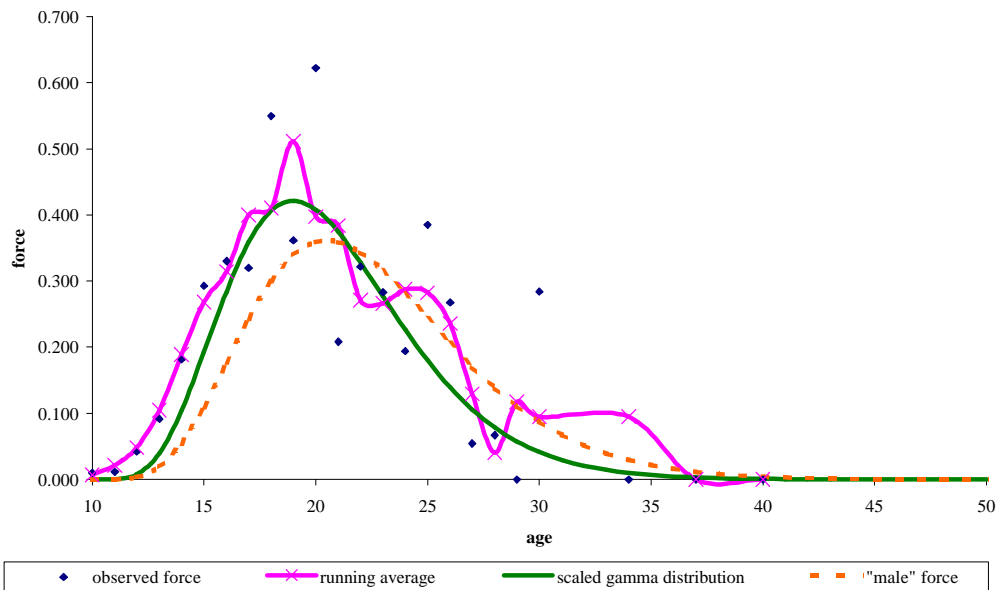


Figure 5

Force of becoming sexually active - Uganda DHS

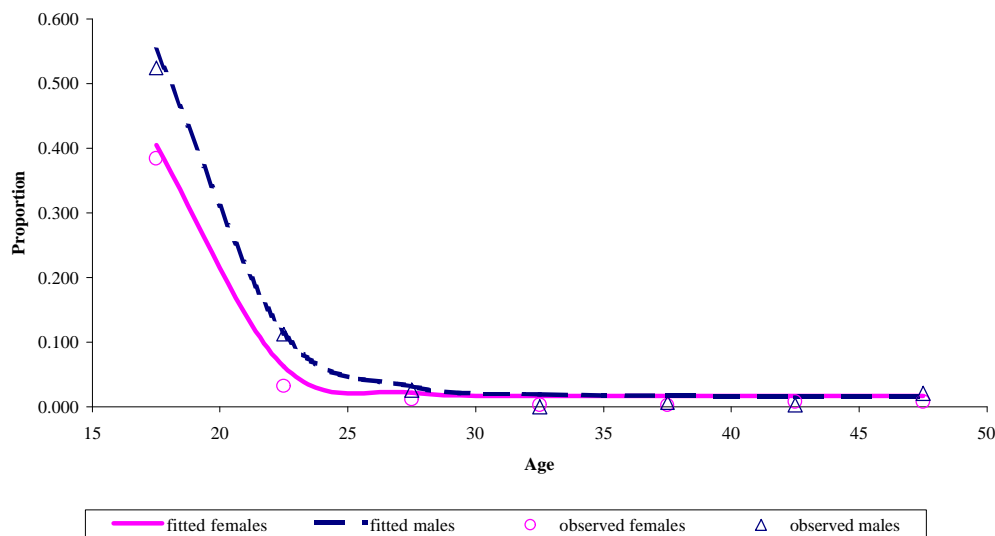


This comparison is shown in figure 4, which also shows a smoothed fitted survivorship function for this lifetable, which was generated using a scaled cumulated gamma function. The curve defining the force of becoming sexually active corresponding to this fit is shown in figure 5, along with the single year forces implied by the raw data on age at first sex. Fitting was done by minimizing the difference between model and observed cumulants rather than model and observed forces, since the former are more robust, and because they correspond to the proportion in the population who have ever had sex, which in itself is an important input into our model. The peak age for the model of the force of entry into sexual activity for women is 19, with about 1.5% of women remaining sexually inactive throughout their lives.

To obtain estimates for males, we assumed that the distribution of age at entry into sexual activity for men would follow the same basic shape as for women, but would occur at a later age. We therefore kept the same basic gamma distribution to describe the rate of entry, but altered the horizontal and vertical scaling to produce an age pattern of proportions who had ever had sex which matched the five-year age data available from the DHS tabulations (Uganda Statistics Department, 1996, p.81). The comparison is shown in figure 6. No minimization routine was used in this case, the fitting was just done by eye.

Figure 6

**Proportions never had sex:
comparison of model and DHS five-year tabulation**



In constructing the model we included a process of exiting from sexual activity at older ages, since “terminal” sexual abstinence is a practice which is found in some of the populations of Eastern and Southern Africa, and since celibacy may well become more prevalent as a reaction to the dangers of forming new sexual relationships, particularly among the divorced and the widowed. However, the empirical data for giving structural form to forces of leaving the sexually active state are much harder to come by. There are many definitional problems, since a person who has been celibate for a long period of time could become sexually active again.

Figure 7

Force of long-term sexual abstinence - Uganda DHS

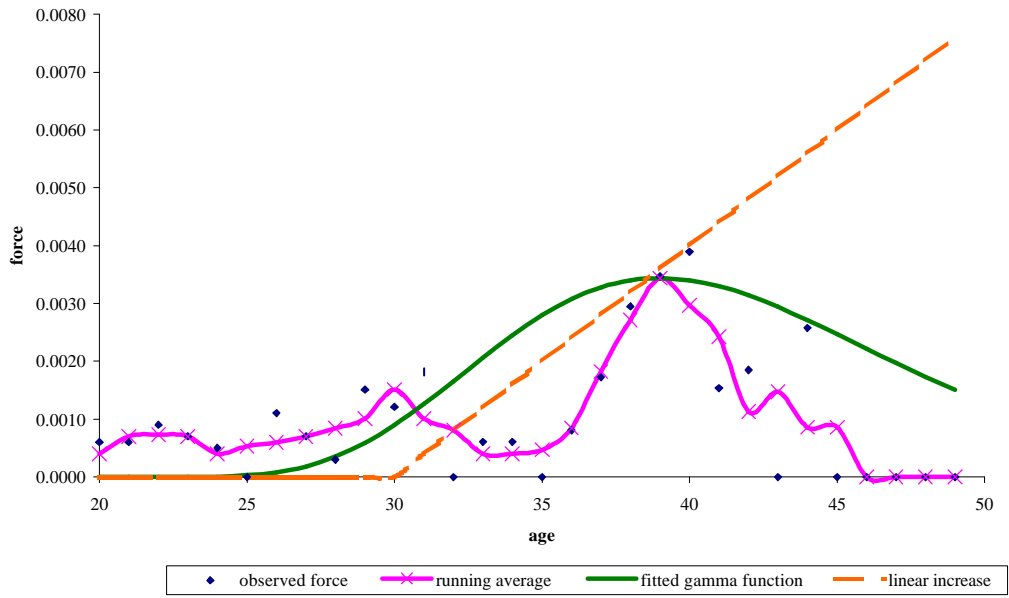
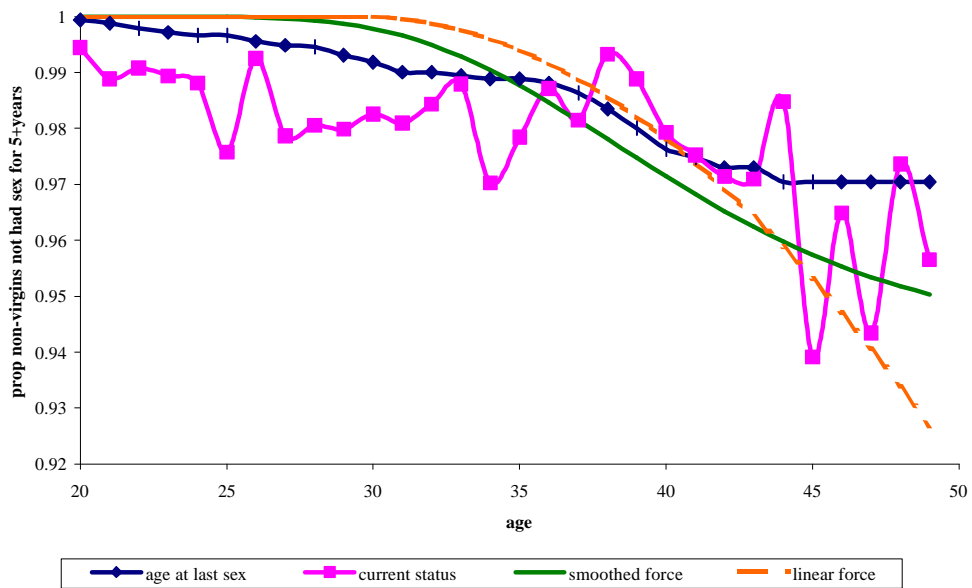


Figure 8

Long term abstinence lifetable - Uganda DHS



We did some investigations using “not had sexual intercourse in the last 5 years” as an index for terminal abstinence, constructing lifetables from individual records for females, on the basis of both current status and reported age at last sex (for those abstaining 5+ years), in the same way as with the data on entry into sexual activity. The results are shown in figures 7 (forces) and 8 (status). Although we attempted fitting a curve of the same kind of functional form as was used to smooth entry into sexual activity, we did not think the data warranted such efforts. It is difficult to discern a pattern, the numbers at risk are small (there are fewer women in the older age groups – women at single years of age in their 40s are numbered in the hundreds, those in their 20s are numbered in thousands) and the events themselves are more rare – peak single year rate of entry into abstinence is of the order of 0.4 %, whereas for entry into sexual activity it is over 50%). However, the cumulated effect of these small forces is not quite negligible – by age 45 about 5% of women report not having had sex for 5 or more years. We therefore decided to use a very simple function to represent the force of becoming celibate. We assumed that for women this would be zero up to age 30, and would then increase at the rate of 0.0004 per year. The force and its effect on proportions celibate are shown as the dashed line in figures 7 and 8.

For males, the decision about the form of this force was even more arbitrary, as there are no equivalent tabulations on long term abstinence for men. We assumed that the celibacy force for men was zero until age 40, and then increased at the rate of 0.0003 per year. The only piece of supporting evidence for making male celibacy in later life a more rare event than for females, is the published statistic which indicates that about 55% of women over 40 reported not having had sex in the last four weeks, whereas for men the equivalent figure was under 30%. This pattern would go with general social observations about partnership structures in Uganda, with a typical age difference of 5 to 10 years between partners, and higher re-marriage rates for widowers than for widows.

Putting together the forces of entry into and exit from sexual activity yields an age pattern of proportions currently sexually active by age shown in figures 9 and 10 for females and males respectively. These proportions active are used to derive age specific fertility rates for sexually active females and to estimate an age pattern of incidence rates for the sexually active, from incidence data which take the whole of the HIV negative population as their denominator.

Numerical values for the forces of entry and exit from sexual activity by single year of age for males and females can be found on the sheets *upsilon* and *tau* in the workbook *UgandaIn.xls*; the resulting male and female proportions sexually active by single year of age are shown in the sheet *YPROP* in the same workbook. The workbook which performs the fitting of the functional form for the force of start of sexual activity to the empirical data on age at first sex (the basis of figures 4 to 8 and the numbers in *upsilon* and *tau*) is not yet in a user-friendly format, and therefore not available yet on the web site.

Figure 9

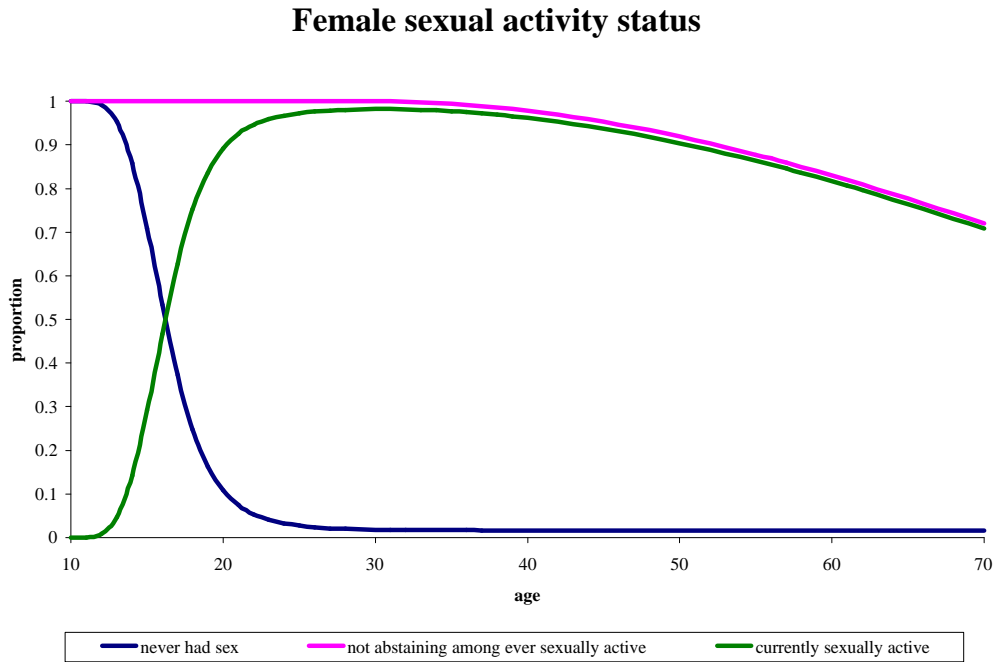
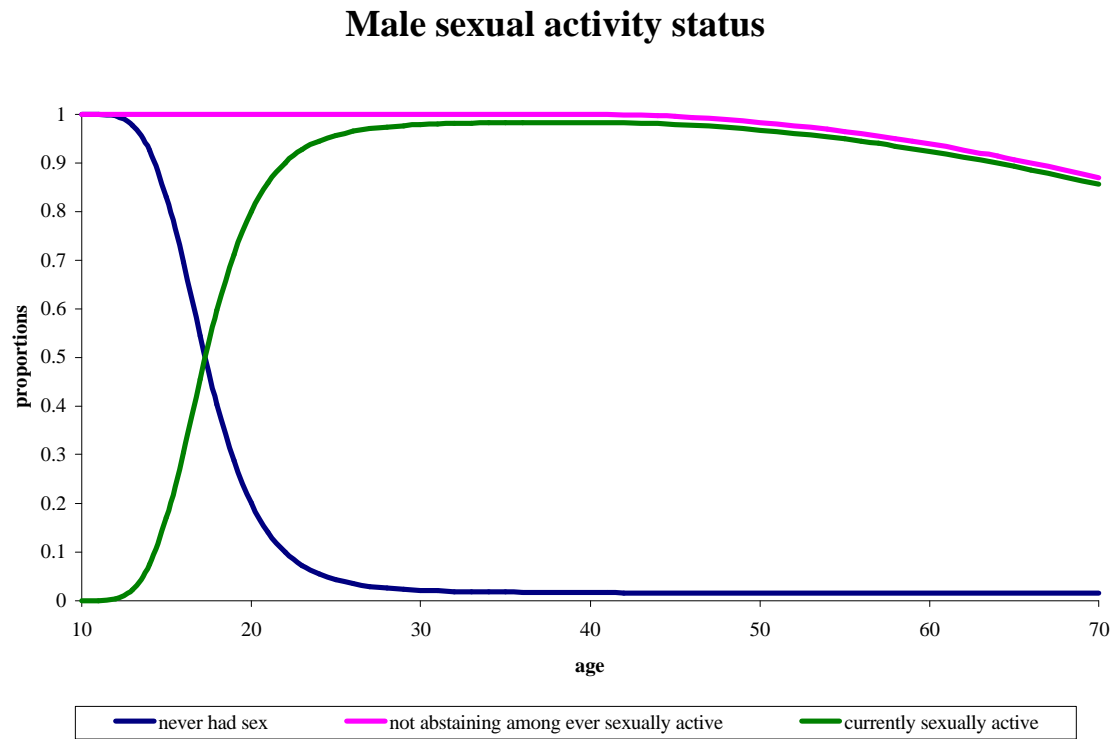


Figure 10



Fertility

Single year age specific fertility rates were calculated from downloaded DHS birth history data, smoothed using running averages, and scaled with reference to published DHS tables (Uganda Statistics Department, 1996, pp. 29-31). To convert the conventional fertility rates to rates for the sexually active these were divided by the proportion sexually active (also derived from analysis of DHS data, see below). At ages under 17, when the proportions sexually active are relatively small, dividing out by this proportion can produce highly irregular and implausible values – at these young ages the computed result was therefore compared with a plausible reference distribution notionally representing biological fecundability at these young ages, which logically cannot be exceeded. The resulting single year data for fertility in the sexually active population are shown in the worksheet *beta* in the workbook *UgandaIn.xls*. Figures 11 and 12 show the two stages of the construction process for this distribution. Note that we have used a fertility distribution incorporating a half-year shift to account for the fact that we use non conventional rates in the projection (see model design chapter, notes following equation 15).

It may well be easier to obtain smooth, well behaved data for the youngest ages by working with single-year fertility models fitted to fertility data tabulated in conventional five year age groups, but we have not yet investigated this with Ugandan data. The workbook showing the fertility calculations has not yet been tidied up enough to make available on the web site.

Figure 11

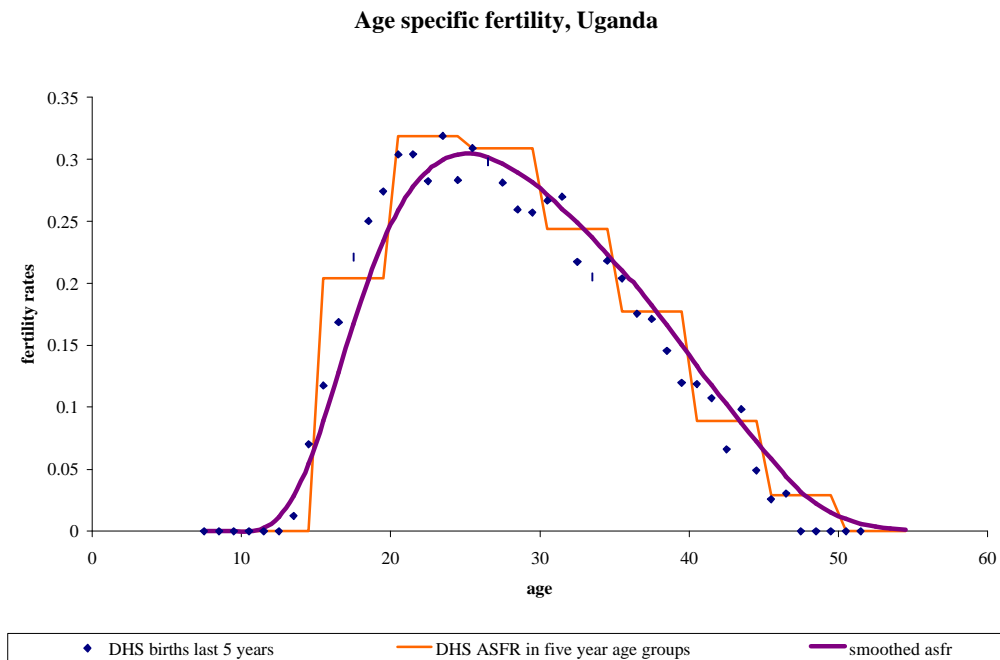
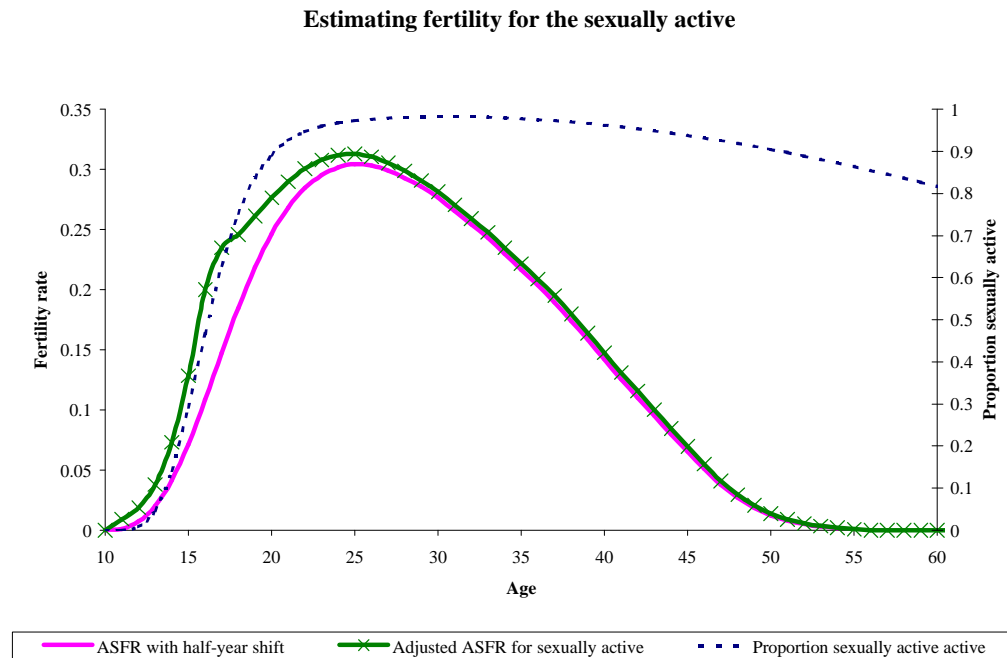


Figure 12



Force of infection

Incidence data are much less common than data on prevalence, as accurate data are only available for populations which are under continuous surveillance, with regular HIV testing and monitoring of demographic events. Data from five such studies are shown in table 2. Two of the studies took place in Uganda (Masaka and Rakai), the other three (Kisesa, Mwanza and Kagera) are from Northern Tanzania, from regions just across Lake Victoria from the Ugandan study sites. All of these studies were located in areas where the epidemic had been established for around five to ten years when data collection took place. The Mwanza study was designed as a test of the possibility of controlling the spread of HIV by providing enhanced detection and treatment services for other STDs which are co-factors in the transmission of the HIV virus. Data from this study are reported separately for the control and intervention areas. Some of the other studies also provided interventions of this kind eventually, but the data reported here are from the pre-intervention phases, when information and education services were the only form of intervention.

Figures 13 and 14 show how the data were amalgamated to produce “standard” incidence schedules for males and females. Basically, the process involved (i) dropping the Rakai study, which had erratic patterns because of the relatively small numbers in this study, (ii) splitting the data from the other studies into conventional five-year age groups, using direct standardization and weighting adjacent age groups to allow for an annual growth rate of 3%, and then (iii) adding the incident cases and person-years of observation from the different studies to obtain combined incidence rates. The final choice of a standardised incidence pattern was based on a combination of data from Masaka, Kisesa, Kagera and the Mwanza control area. The data from the Mwanza intervention area were not used because of their atypically flat age pattern. This yields a combined total of 14,800 PYO for males and 16,600 PYO for females. The workbook showing the detailed calculations for producing this standard is called *incbyage.xls* – it is not yet fully documented, but is largely self explanatory.

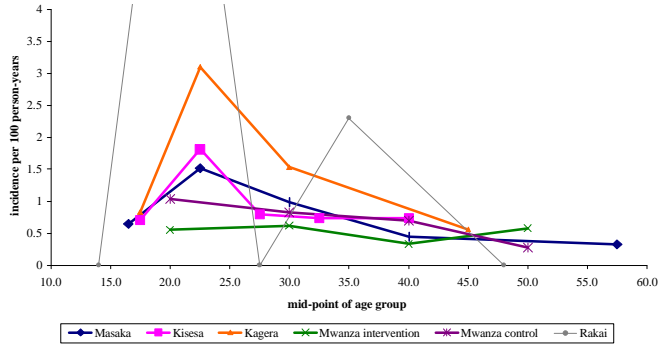
Table 2 Incidence data from East African studies

Years	Location	start age	end age	male incidence per 100 PYO	male person-years observed	female incidence per 100 PYO	female person-years observed
90-94	Masaka	13	19	0.21	1934	0.65	1849
		20	24	1.03	678	1.52	658
		25	34	1.25	958	0.99	1309
		35	44	1.16	688	0.45	887
		45	+	0.72	1813	0.33	1813
		13	+	0.72	6071	0.69	6517
94-97	Kisesa	15	19	0.20	994	0.71	706
		20	24	0.76	788	1.81	846
		25	29	0.90	553	0.80	746
		30	34	1.31	469	0.74	678
		35	44	0.93	750	0.74	1089
		15	44	0.73	3544	0.84	4065
87-89	Kagera				pyo 15-24		pyo 15-24
		15	19	0.00		0.83	
		20	24	0.14	292	3.10	378
		25	34	3.09	259	1.53	522
		35	54	0.72	414	0.56	540
	15	54	1.35	965	1.39	1440	
91-94	Mwanza intervention area	15	24	0.53	1322	0.55	1442
		25	34	0.54	1120	0.62	1292
		35	44	0.79	764	0.33	898
		45	54	0.78	768	0.58	692
		15	54	0.63	3974	0.53	4324
91-94	Mwanza control area	15	24	0.86	1520	1.04	1634
		25	34	1.23	1222	0.83	1332
		35	44	1.33	752	0.69	866
		45	54	1.06	754	0.28	720
		15	54	1.08	4248	0.79	4552
89-90	Rakai	13	14	0.0	45	0.0	40
		15	19	0.0	70	5.0	81
		20	24	9.2	54	6.8	73
		25	29	2.5	41	0.0	75
		30	39	1.4	71	2.3	88
		40	55	1.3	159	0.0	131
		13	55	2.1	439	2.3	488

need to get refs to published sources from Elizabeth P.

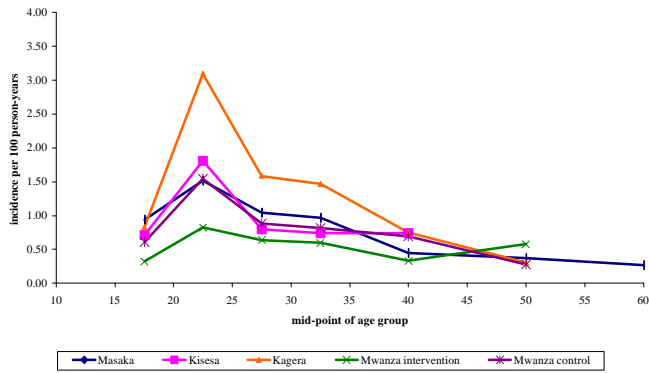
Fig 13
(a)

Female age specific HIV incidence rates: published data



(b)

Female age specific HIV incidence - data split into standard age groups



(c)

Female age specific HIV incidence rates: combined data

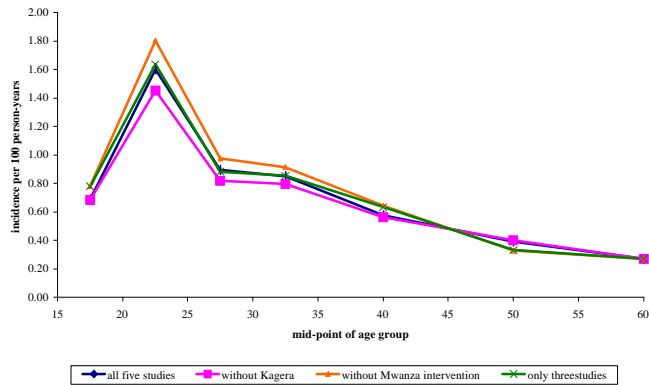
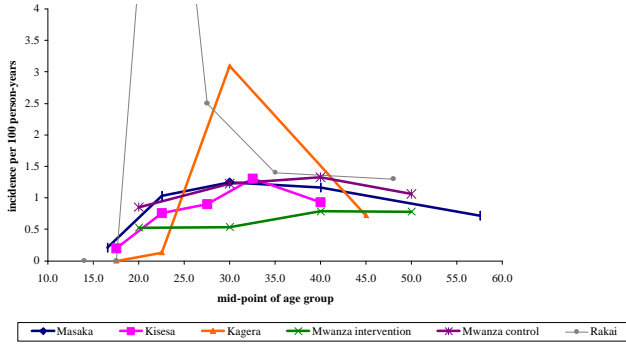


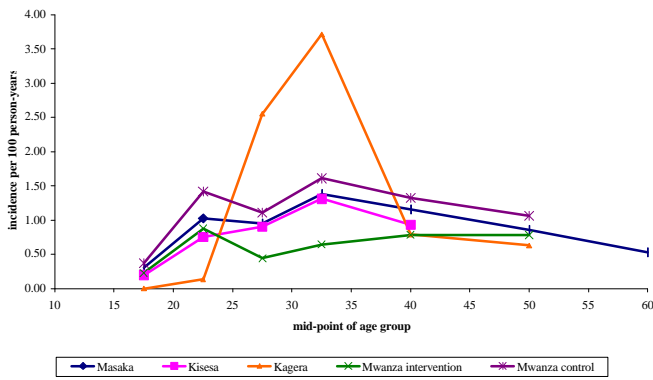
Fig 14
(a)

Male age specific HIV incidence rates: published data



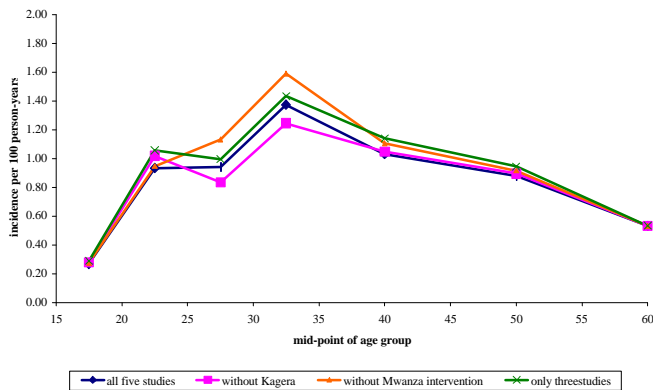
(b)

Male age specific HIV incidence - data split into standard age groups



(c)

Male age specific HIV incidence rates: combined data



The standard age specific incidence patterns for males and females were converted to patterns applicable to the sexually active by dividing through by the smoothed proportions sexually active in each age group obtained earlier. This calculation is shown in table 3.

Table 3 Estimation of standard incidence patterns for sexually active

age group	standard % incidence rates from combined longitudinal studies		smoothed proportion sexually active		% incidence rates for sexually active	
	males	females	males	females	males	females
15-19	0.275	0.783	0.596	0.752	0.461	1.041
20-24	0.948	1.804	0.927	0.959	1.022	1.881
25-29	1.132	0.977	0.974	0.980	1.162	0.996
30-34	1.588	0.916	0.982	0.981	1.617	0.933
35-44	1.105	0.647	0.984	0.962	1.123	0.672
45-54	0.915	0.330	0.968	0.905	0.945	0.365
55-64	0.533	0.273	0.924	0.817	0.577	0.334

Figure 15

Smoothing female incidence rates for the sexually active

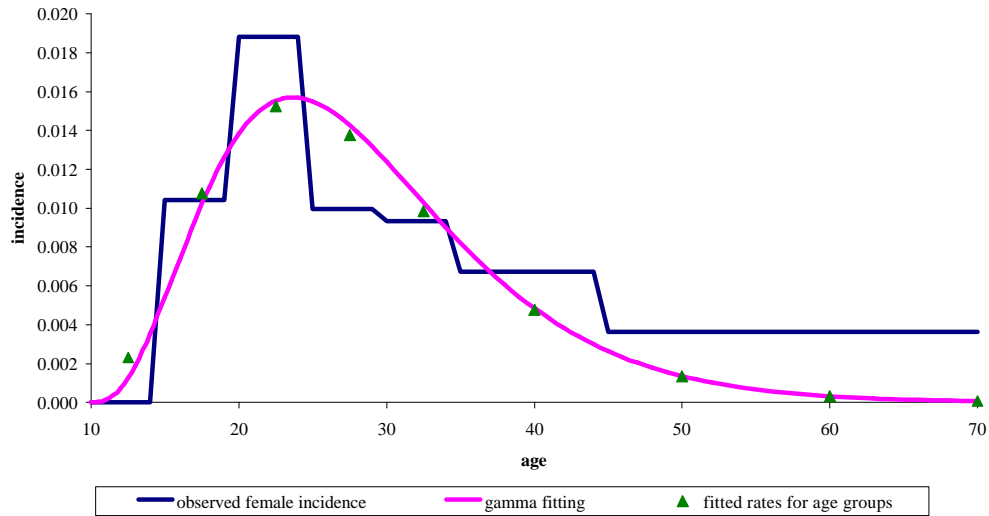
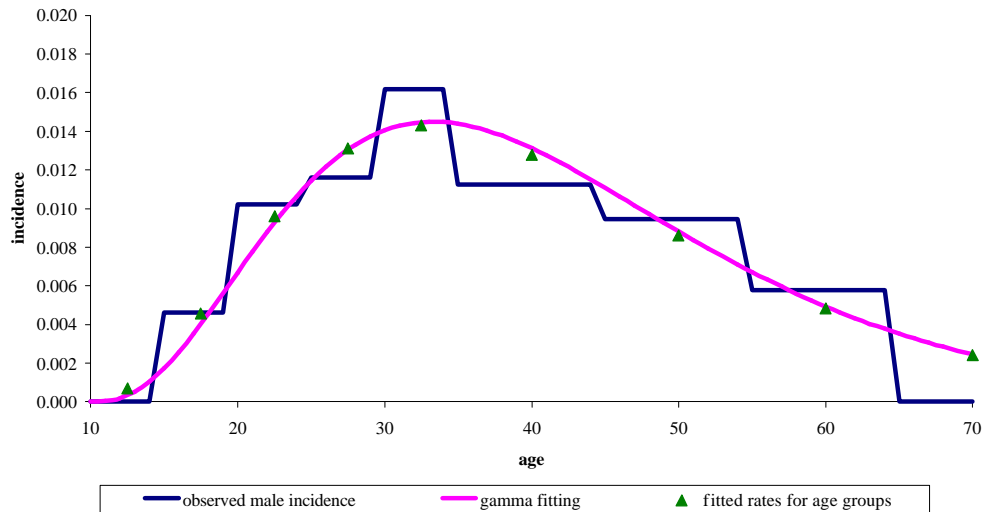


Figure 16

Smoothing male incidence rates for the sexually active



Figures 15 and 16 show how the incidence rates for the sexually active in the broad age groups shown in table 3 were fitted using a force of infection function in the shape of a gamma distribution, with vertical and horizontal scales chosen to minimize the differences between observed and model incidence rates in each age group. The fittings produced a peak incidence rate for sexually active females of around 1.6% at age 24, and for males the peak rate of 1.4% is reached at around age 34. These smoothed, single year distributions are normalized and used as inputs for the force of infection age pattern. The normalized shape can be scaled over time to represent increasing or declining incidence trends.

The worksheet *phi* in the workbook *UgandaIn.xls* shows the single year values for the age-specific incidence pattern used in the projection. The workbook *forcephi.xls* shows how the smoothing was achieved, though this is not yet properly documented.

Coefficients for level of force of infection

The aim of the projection is to fit prevalence predicted by the model to observed data, and then make useful and realistic extrapolations for the short to medium term. The procedure which we adopt is therefore to use data for Uganda and closely related populations to fix mortality, fertility and sexual behaviour parameters described above. In this first test of the model we have assumed that fertility, mortality and age at start and end of sexual activity will not change over time, although these assumptions could be altered on the basis of empirical evidence if any were available. The age pattern (but not the level) of the force of infection has also been fixed, as described above.

However, it would be extremely difficult to find empirical evidence to make direct estimates of the three coefficients which govern the time trend in the level of the force of infection. In the model description chapter these are referred to as iota (i = the component independent of prevalence), kappa (k = the constant multiplier for prevalence) and theta (q = the time varying

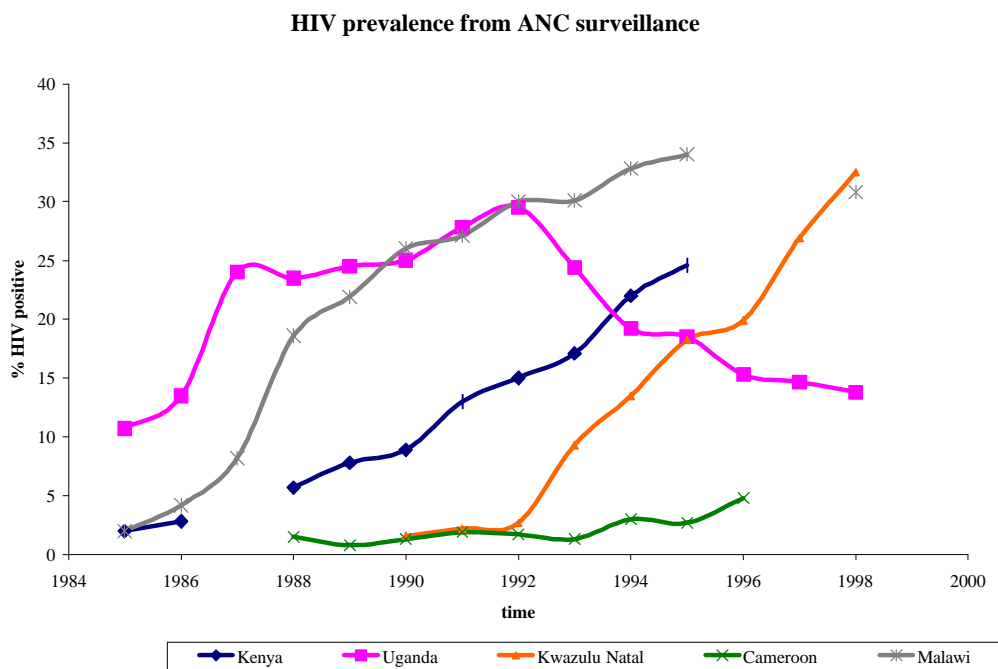
prevalence multiplier). We therefore propose that this parameter evaluation is done indirectly, using a trial and error procedure (or some kind of search routine) to locate suitable values for these parameters that generate a reasonable fit to the observed time sequence of HIV prevalence values in the population. Since the “bare bones” model (described in a previous chapter) uses similar ι , κ and θ parameters, but is relatively straightforward to fit, the parameter values generated from the fitting of this non-age structured model may be used as starting points for fitting the full model.

Table 4 – % prevalence of HIV in pregnant women in Uganda, 1985-98

1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
10.7	13.5	24.0	23.5	24.5	25.0	27.8	29.5	24.4	19.2	18.5	15.3	14.7	13.8

Table 4 shows the time series of prevalence values used to determine the components of the force of infection. The time trends for prevalence in Uganda are compared with those from other African countries in figure 17. We can see that Uganda is quite unusual in recording a steady decline in prevalence in recent years. Because it has this distinct sharp peak around 1992, we might expect that a close-fitting model would be relatively hard to find, but that if a good fit existed it would have reasonably strong predictive power, at least in the short run, because of its uniqueness. On the other hand, the very sharp peak may be an artefact of the data, as these are based on measurements in pregnant women. When age at entry into sexual activity increases, prevalence measures in pregnant women tend to exaggerate the pace of change in prevalence in the general population.

Figure 17



The “bare bones” model is capable of producing a reasonable fit to the Ugandan data, as is shown in figure 18, and suggests that long-term endemic prevalence in Uganda might be heading for a value round 2.5%, and that reasonable starting values in our search for force of infection parameters would be 0.003 for ι , (the prevalence independent component); κ (the constant prevalence dependant element) might be around 0.03; and the time dependent prevalence multiplier which is allowed to fall to zero fairly soon after the start of the epidemic would need to be much larger, at around 0.55. The “bare bones” fitting also suggests that the notional “zero” year for Uganda was 1977, and that rapid behavioral change began in 1986, and was completed by 1990. These timing values could also be useful in trying to locate a close fitting “minimalist model” to describe the Ugandan epidemic.

Figure 18

