Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa

Basia W. Zaba, Lucy M. Carpenter[a,b], J. Ties Boerma[c], Simon Gregson[d], Jessica Nakiyingi[b] and Mark Urassa[e]

Objectives: To find a simple and robust method for adjusting ante-natal clinic data on HIV prevalence to represent prevalence in the general female population in the same age range, allowing for fertility differences by HIV status.

Background: HIV prevalence comparisons for pregnant women and women in the general community show that prevalence in the latter is significantly higher than in the former. An adjustment procedure is needed that is specific for the demographic and epidemiological circumstances of a particular population, making maximum use of data that can easily be collected in ante-natal clinics or are widely available from secondary sources.

Methods: Birth interval length data are used to allow for subfertility among HIV-positive women. To allow for infertility, relative HIV prevalence ratios for fertile and infertile women obtained in community surveys in populations with similar levels of contraception use are applied to demographic survey data that describe the structure of the population not at risk of child-bearing.

Results: For populations with low contraception use, the procedure yields estimates of general female HIV prevalence of 35–65% higher than the observed ante-natal prevalence, depending on population structure. Results were verified using general population prevalence data collected in Kisesa (Tanzania) and Masaka (Uganda). For high contraception use populations, adjusted values range from 15% higher to 5% lower, but only limited verification has been possible so far.

Conclusions: The procedure is suitable for estimating general female HIV prevalence in low contraception use populations, but the high contraception variant needs further testing before it can be applied widely.

AIDS 2000, **14**:2741–2750

Keywords: demographic estimation, HIV prevalence, measurement bias, population-based surveys, sentinel surveillance

Introduction

Prevalence data are vital for tracking the progress of the HIV epidemic, planning the allocation of resources to care for AIDS patients, and monitoring the impact of interventions. Ideally, these should come from community-based serosurveys but only few of those have been carried out for financial, logistical or ethical reasons. Furthermore, surveys provide only occasional estimates of point prevalence. Because of this, many countries rely on sentinel surveillance systems where easily accessible populations are monitored anonymously over time. The most widely used sentinel group is pregnant women [1]; this source is particularly appropriate for

From the Centre for Population Studies, London School of Hygiene and Tropical Medicine, London, the [a]Department of Public Health, Institute of Health Science, University of Oxford, UK and [b]Medical Research Council Programme on AIDS in Uganda, Uganda Virus Research Institute, Entebbe, Uganda, the [c]Carolina Population Centre, School of Public Health, University of Northern Carolina at Chapel Hill, North Carolina, USA, the [d]Department of Zoology, University of Oxford, UK and the [e]Tanzania–Netherlands Project to support HIV/AIDS control, Mwanza, Tanzania.

Requests for reprints to: B. Zaba, Centre for Population Studies, London School of Hygiene and Tropical Medicine, 49–51 Bedford Square, London WC1 3DP, UK.

Received: 18 April 2000; revised: 23 August 2000; accepted: 4 September 2000.

sub-Saharan Africa, where in many countries, more than 90% of women attend ante-natal clinics [2–7]. In countries where contraception use is low, pregnant women are assumed to be reasonably representative of the sexually active female population. Data from these surveillance systems are, however, potentially subject to three kinds of bias: the purposive selection of clinics for participation in the surveillance program; the self-selection of women who attend them; and the structural bias associated with testing only those women in the general population who become pregnant. The first two of these biases are dealt with by surveillance system designs which aim to ensure that the geographical spread and type of facilities included in the sample correspond with those used by the population under study. The third, however, is more likely to require demographic adjustments and is the subject of this paper.

Research from community-based longitudinal studies shows that HIV-positive women have lower fertility than those who are not infected [8–10]. Similar findings are reported in follow-up studies of hospital cohorts which monitor groups of women whose HIV status is measured at delivery [11–13]. A recent meta-analysis confirms that HIV-positive women have a higher incidence of still births and foetal loss than women who are not infected [14]. However pre-existing fertility differences (e.g. prior infection with sterilizing sexually transmitted infections) may also be important [15]. Other possible reasons for the fertility differential are behavioural and biological, such as higher rates of widowhood and divorce and increased problems with conception among HIV-positive women [16]. A serious consequence of lower fertility in HIV-positive women is that prevalence estimates based on data from ante-natal clinics will tend to be lower than in the general population [17,18]. This is generally the case for all but the youngest age groups where selection effects predominate, as virgins are almost all HIV-negative. Estimates of HIV prevalence from a range of African ante-natal clinic data for women aged 15–44 years were found to be on average 28% lower than those for all women in the general population [19].

Various approaches have been proposed for improving the estimation of community-based prevalence from ante-natal clinic data. Although age standardization has been suggested [20], this actually increases the underestimate of crude prevalence for women 15–49 years because it gives more weight to the oldest age groups where prevalence differences are largest. For an adjustment method to be widely applicable it needs to yield robust estimates, be sensitive to local fertility and epidemic circumstances, be simple to compute, and not require too much extra data collection. One possible approach is a direct comparison between fertility rates for women who know they are HIV positive and those for the population as a whole [21–24]. While potentially workable in very low prevalence populations, this approach has yet to be tried in African countries where very few infected women know their HIV status. Other approaches involve mathematical models such as those developed to explore the scale of the prevalence differential, its sensitivity to behavioural and biological determinants, and its trend over time in an evolving epidemic [25,26]. While these models have yet to be refined and tested, future developments may allow prediction of age-specific prevalence among women in the general population on the basis of prevalence in pregnant women of the same age.

Although conventional age-specific fertility rates cannot be obtained directly from ante-natal clinic data, it is possible to get other important fertility indicators from this source. The distribution of births by order can yield an indirect estimate of fertility level [27], and birth intervals furnish information on the current tempo of fertility. These indirect methods of fertility estimation cannot be adapted easily to obtain differential age-specific fertility rates by HIV status, as they usually assume unchanging fertility patterns. Moreover, parity levels of the HIV-positive women will reflect in part their fertility prior to infection. However, parity and birth interval data collected in ante-natal clinics can be used for the purposes of adjusting ante-natal clinic-based HIV prevalence estimates [28]. The purpose of this paper is to describe a recently developed adjustment method and to illustrate it with application to data from Uganda, Tanzania and Zimbabwe.

Materials, methods and assumptions

The proposed approach divides the problem into two parts. First, using information on birth intervals obtained from ante-natal clinics it allows for subfertility among those HIV-infected women who are still capable of bearing children. Secondly, it allows for different levels of HIV prevalence among women who are not seen in ante-natal clinics, by using data from demographic surveys to ascertain what proportions of the population are not at risk of childbearing, because they have effectively become sterile, or are using contraception or are not sexually active. HIV prevalence ratios for these population subgroups relative to those women who are at risk of childbearing can be estimated from community-based cohort studies.

Table 1 is a schematic guide to the proposed breakdown of the female population according to fertility category. Women are divided into mothers and those who have never borne a child. The adjustment procedure is carried out separately for these two groups for

Table 1. Definitions of categories used in the proposed adjustment method.

Fertility category	Childless women	Mothers
Never had sex	Mainly teenagers, zero HIV prevalence	
Had sex, but not currently sexually active	Had first sexual encounter but not cohabiting	Widows, separated, divorced, abstaining
Sexually active, fecund, users of contraception	Using contraception to delay first birth	Using contraception for spacing or limitation
Sexually active, infecund	Suffering from primary infertility	Suffering from secondary infertility
Sexually active, fecund, potentially at risk of conception[a]	Women in early casual and cohabiting unions	Women with recent birth
Pregnant[a]	Expecting first child	Expecting higher order birth

[a]These categories are combined as 'sexually active, fecund' in the following tables.

two reasons. Firstly, it is possible to allow for subfertility only in the former group, for whom measurements can be made of the past tempo of childbearing; secondly, the two groups have a very different risk composition with respect to HIV infection. Childless women are composed in a large part of women who have never been sexually active or have started sexual activity relatively recently. Their HIV prevalence levels are therefore considerably lower than those of women who have been sexually active for some time. Amongst sexually active childless women, the fecund are defined as those who are expecting their first child, or who experienced first sexual activity less than 5 years ago and are not using contraception. Fecund mothers are defined as those who are currently pregnant with a higher order birth, or who have had a birth in the last 5 years and are sexually active and not using contraception. 'Sexually active' is defined as having had sex in the last year. The 5 year cut-off point for births is somewhat arbitrary, but is often used in the operational definition of infertility [29].

In practice, some of the groups in Table 1 are likely to be much more important than others, either because of their numerical size, or because of their distinctive HIV risk profile. For example, those who have never had sex – a sizeable proportion of childless women because of the numerical dominance of the youngest age groups in growing populations – clearly have zero (or almost zero) HIV prevalence. By contrast, those suffering from secondary or primary sterility would contain a large proportion of women with a history of serious infections with other sexually transmitted infections, and although not numerically dominant, this group will tend to have high HIV prevalence compared with those who are fecund.

Childless users of contraception are likely to be a very small group in sub-Saharan Africa, where modern family planning methods are used mainly for child spacing and for family limitation by women who have already had several children. However these users deserve particular attention, as they are likely to include persons whose sexual behaviour is atypical. They may be using contraception just to prevent childbearing, and this would tend to place them among the better educated who might be expected to avoid casual unions with risky partners – although their increased mobility and wealth may also mean that they (or their partners) are at higher risk of forming extra-marital relationships. On the other hand, effective users of condoms are also included among contraception users not seen in ante-natal clinics: those who have always used condoms prophylactically should be at low risk, but condom use is often initiated in a partnership where one or other partner becomes aware that they are already infected; relatively high HIV prevalence would be expected in this subgroup.

The most fertile women – those with regular menstrual and ovulatory cycles, without tubal occlusions, in cohabiting unions with high coital frequency rates – will on average experience shorter birth intervals than subfertile women. The probability that in a particular year a mother reaches the later stages of pregnancy at which ante-natal check-ups are usually performed can be approximated by the inverse of her birth interval, measured in years. Over-representation of the most fertile women in ante-natal clinics is therefore compensated for by weighting mothers' reports by their birth interval. Women who happen to be pregnant at a given time (bottom right cell of Table 1) constitute about 6–11% of the female population aged 15–49 years in sub-Saharan African countries. This weighting makes them representative of the population who are potentially able to experience pregnancy if they waited long enough (the last two cells in the right hand column). If 'recent' is defined as 'having had a birth in the last 5 years' this accounts for between 20 and 45% of all women in these populations.

In theory, it would be possible to weight reports for the primigravida in a similar way, by finding out how long they had been in a sexual relationship with the partner whose child they were currently expecting. However, this is not a question which is routinely asked in an ante-natal clinic setting, and such information might not be well reported in any case, as the precise beginning of a relationship is not always easy to date, and may not be known at all in the case of a birth which was the result of a casual relationship. Furthermore, women who had recently been married would

tend to report the marriage date as the start of the relationship, even if they had been sexually involved with their partner prior to this. In a population with a very narrow distribution of age at first sex, and a rapid progression from first sex to cohabitation, it might be possible to obtain estimates of relative waiting time to first birth from the ages of the women pregnant for the first time. However, in most populations estimating waiting time to first birth from age alone would risk giving too much weight to the experience of women who simply started sexual activity at a later age. Consequently, data for primigravida are used unweighted, analyses are performed separately for mothers and childless women, and the two groups are merged at the end according to their relative size in the population as a whole.

The methods described above are illustrated using data from three African studies: the Kisesa study, conducted by the TANESA/National Institute for Medical Research project in the Mwanza region in Tanzania [30]; the MRC General Population Cohort study in Masaka, Uganda [8]; and the Manicaland HIV/STD Prevention Project in Eastern Zimbabwe [31]. Data from these studies are used to estimate HIV prevalence according to childbearing risk category for two 'standard' populations. The low contraception standard was obtained by pooling data from the Tanzanian and Ugandan studies, where contraception use prevalence is around 5%, and should be used for adjustment of ante-natal clinic data in settings where reported contraception use is less than 15%. Data from the Manicaland study, based on a population with nearly 50% contraceptive prevalence, were used to estimate standard relative rates for populations reporting contraception use of 20% or more.

Although prevalence data are routinely collected in ante-natal clinics in many African countries, the only data generally passed on with the anonymous blood sample when it is sent for testing is the mother's age. The suggested adjustment procedure requires that the sample is also categorized according to whether it was taken from a woman who is pregnant for the first time, and for those who are expecting higher order births, the birth interval is needed. Although data on birth interval and parity are routinely collected in ante-natal clinics, they are not yet widely reported in tables pertaining to HIV prevalence. We therefore test our procedure using data for pregnant women collected in the MRC General Population Cohort and the Kisesa study. As direct estimates of prevalence are also available for these cohorts, it is possible to verify how closely the resulting adjustment approximates true prevalence in the general female population.

Results

Estimation of relative risk for standard populations

For the low contraception standard, HIV prevalence estimates ranged from 0.6% among women who had never had sex to 19.5% among sexually active infecund childless women (Table 2). Data presented here refer to the age group 15–49 years, but the same method could be adapted for narrower age ranges in populations in which childbearing is more concentrated. Prevalence estimates were generally higher for the high contraception standard, ranging from 2.1% for women who had never had sex to 52% among sexually active, infecund, childless women. Relative prevalence levels

Table 2. HIV prevalence by childbearing risk category for low and high contraception standards as estimated from community-based cohort studies.

	Low contraception standard[a]		High contraception standard[b]	
Fertility category	HIV prevalence % (n)	Relative prevalence	HIV prevalence % (n)	Relative prevalence
Childless women	(1197)		(1354)	
Never had sex	0.6 (510)	0.10	2.1 (933)	0.09
Had sex, not sexually active	1.6 (129)	0.27	20.9 (67)	0.92
Sexually active, using contraception	14.3 (35)	2.48	22.0 (41)	0.96
Sexually active, infecund	19.5 (159)	3.38	52.0 (50)	2.28
Sexually active, fecund (reference category for childless women)	5.8 (364)	1.00	22.8 (263)	1.00
Mothers	(4206)		(3775)	
Not currently sexually active	16.4 (256)	2.09	41.5 (615)	1.49
Sexually active, using contraception	15.6 (243)	1.99	26.4 (1630)	0.95
Sexually active, infecund	12.5 (634)	1.59	39.8 (595)	1.43
Sexually active, fecund (reference category for mothers)	7.8 (3073)	1.00	27.8 (935)	1.00
Total	(5373)		(5129)	

[a]Estimated from pooled data collected in Kisesa (1994–1998) [30] and Masaka (1995–1996) [8]. [b]Estimated from data collected in Manicaland (1999) [31].

are also shown in Table 2, separately for first births and higher-order births – women with recent births constitute the reference category in each case. Relative prevalence values reveal sharp differences in infection patterns between the low and high contraception standards. In the low contraception standard, fecund women are a low risk group compared to others who are sexually active, whereas in the high contraception standard the risk level is closer to that of other sexually active women. In the high contraception standard, users have a marginally lower HIV prevalence than those exposed to the risk of childbearing, whereas in the low contraception use standard the prevalence is about twice as high as those we presume are fertile. Infecund women have higher prevalence levels than sexually active fecund women in both standards, but differentials are sharper in the low contraception standard. In this standard there is a reasonably large group of childless women who have had sex at some time in their lives, but are no longer sexually active. These women have very low prevalence, whereas in the high contraception standard this group is numerically much smaller and has a prevalence level similar to childless users of contraception.

Application of adjustment procedure to Ugandan and Tanzanian data

A test application of the proposed method of adjustment in the MRC General Population Cohort is shown in Table 3. Pregnancy-based HIV prevalence rates shown in column (a) were obtained by re-analysing published fertility data from seven annual survey rounds of this cohort to produce new classifications of HIV status by parity. The HIV prevalence rate for mothers shown in column (b) is estimated from individual records by weighting data for mothers experiencing higher order births by their preceding birth interval. Relative prevalence levels for the low contraception standard, shown in column (c) are most appropriate for adjusting data from Uganda and are used to estimate the expected prevalence according to risk category [column (d)]. The population distribution for Ugandan women by fertility category [column (e)] is obtained from a special tabulation of the national Demographic and Health Survey (DHS) data. Expected numbers infected [column (f)] are obtained by applying the expected prevalence to the distribution of women from the DHS. These are then summed to obtain estimated prevalence rates in the population [column (g)].

The best estimate of prevalence obtained using this procedure is 10.3%. This falls within the range of community-based estimates obtained at the beginning (round 2 = 10.1%) and end (round 7 = 10.9%) of the birth monitoring interval. The adjusted estimate represents an increase of about 35% in the observed prevalence of 7.7% in pregnant women. Had birth

Table 3. Adjustment of parity-based HIV prevalence estimates. Worked example using data from the MRC General Population Cohort in rural Uganda, 1989–1996 [8].

Fertility category	HIV prevalence in women giving birth		Relative HIV prevalence for low contraception standard (c)	Expected HIV prevalence in fertility category (d) = (c)(b)	Percentage distribution of women Uganda DHS [2] (e)	Expected HIV infected in fertility category (f) = (d)(e)/100	Estimated HIV prevalence rates in population (g) sums of (f)
	Observed (a)	Weighted (b)					
Childless women					(22.3)		1.51/22.3 = 6.8
Never had sex			0.10	0.82	10.5	0.08	
Had sex, not sexually active			0.27	2.17	1.8	0.04	
Sexually active, using contraception			2.48	19.96	1.3	0.26	
Sexually active, infecund			3.38	27.24	2.2	0.60	
Sexually active, fecund	8.06	8.06	1.00	8.06	6.5	0.52	
Mothers					(77.7)		8.74/77.7 = 11.3
Not currently sexually active			2.09	17.05	14.0	2.38	
Sexually active, using contraception			1.99	16.25	9.1	1.48	
Sexually active, infecund			1.59	12.95	9.0	1.17	
Sexually active, fecund	7.52	8.15	1.00	8.15	45.6	3.72	
All women	7.66				100.0		1.51 + 8.74 = 10.25

DHS, Demographic and Health Survey.

interval length not been adjusted for a slightly lower prevalence estimate of 9.5% would have been obtained.

As a further check on the validity of the method, it was applied to prevalence estimates based on pregnancies in the Kisesa study. The pregnancy-based HIV prevalence estimate obtained from births occurring between the two sero-survey rounds in Kisesa was 4.6% (5.3% after weighting by birth interval) compared with community-based estimates ranging between 6.7% and 7.9% at the 1994 and 1997 survey rounds (average 7.3%). With the population structure of the whole of Tanzania, based on the 1996 DHS, the adjustment procedure yielded an estimate of 6.9% – an increase of 50% over the unadjusted rate. Without weighting for birth interval the adjusted result would have been 6.1%. Using the actual structure of the Kisesa population, the adjusted result was 7.2%, virtually the same as the average community-based estimate.

The adjustment method described above is sensitive to the choice of relative prevalence distributions and population structures. The upper section of Table 4 shows population distributions according to childbearing risk categories for Tanzania, Uganda and four additional Eastern and Southern African countries [2–7]. These distributions have been derived by tabulating individual level data in the DHS country basic recode files, according to the categories described in Table 1. The largest contrasts are found in the proportions who have never had sex, and in mothers who are current users of contraception or who are sexually inactive. Because these groups can be at the extremes of HIV prevalence levels, as suggested by the relative prevalence estimates from the low contraception use standard, this could have a potentially large impact on the adjusted prevalence level.

Adjustment factors shown in the lower section of Table 4 for the low and high contraception use standards are derived from relative prevalence estimates, as defined in Table 2, combined with the DHS population distributions given in the upper section. These suggest that the contrast between relative prevalence in the low and high contraception use standards is more important than the variability between population structures. In particular the adjustment factor for multigravida is much lower when the high contraception standard is used, hardly exceeding 10% in most combinations. As a result, using the high contraception use risk pattern would tend to produce an adjusted prevalence rate almost the same or only slightly higher than that observed in pregnant women whatever population structure is used. It could even suggest an overall downward adjustment in populations dominated by young, childless women. In contrast, the low contraception risk pattern consistently produces adjusted prevalence estimates some 50% higher, as shown in Fig. 1.

In summary, Table 4 suggests that for low contraception use populations (use of less than 15%) female prevalence in the general population is 80% of the prevalence rate in primigravida applied to childless women, combined with 150% of the prevalence rate for multigravida applied to mothers. The low contraception standard has been verified at least for the two study populations (of combined size 5370) on which it was based, and is based on populations in which the epidemic has more-or-less stabilized.

Table 4. Population distribution (%) by childbearing risk categories and adjustment factors for selected Demographic and Health Survey studies.

	Tanzania 1996 [5]	Uganda 1996 [2]	Zambia 1996 [6]	Botswana 1988 [3]	Kenya 1993 [4]	Zimbabwe 1994 [7]
Childless women (subtotal)	(24.7)	(22.3)	(25.3)	(24.6)	(27.7)	(28.3)
Never had sex	13.2	10.5	11.4	7.7	15.7	19.7
Had sex, not sexually active	1.8	1.8	1.9	0.5	1.5	0.7
Sexually active, using contraception	0.6	1.3	1.1	3.8	0.6	0.8
Sexually active, infecund	2.5	2.2	2.3	2.7	2.0	1.2
Sexually active, fecund	6.6	6.5	8.6	9.9	7.9	5.9
Mothers (subtotal)	(75.3)	(77.7)	(74.7)	(75.4)	(72.1)	(71.6)
Not currently sexually active	17.2	14.0	17.4	14.4	8.1	8.2
Sexually active, using contraception	12.1	9.1	9.5	28.3	20.1	29.5
Sexually active, infecund	12.5	9.0	11.4	12.5	10.0	11.8
Sexually active, fecund	33.5	45.6	36.4	20.3	33.9	22.1
Adjustment factors for populations and relative prevalence combinations						
Low contraception standard						
Parity zero	0.7	0.8	0.8	1.2	0.6	0.5
Higher order	1.5	1.4	1.5	1.7	1.5	1.6
High contraception standard						
Parity zero	0.6	0.7	0.7	0.8	0.6	0.4
Higher order	1.2	1.1	1.2	1.1	1.1	1.1

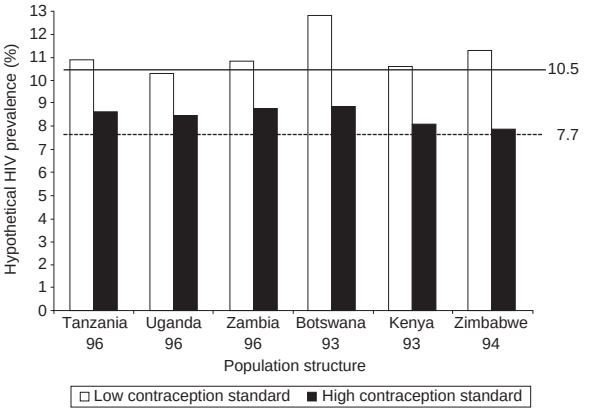

Fig. 1. Prevalence adjustment. Relative effects of population structure and relative prevalence ratios, applied to observed HIV prevalence in pregnant women in Masaka (1989–1996) [8]. The horizontal lines show the unadjusted prevalence level for pregnant women (7.7%) and the prevalence level observed in the community (10.5%) in the same time period. Population structures based on DHS reports [2–7].

In populations with high (20% or more) use of contraception, an adjustment factor of 60% would be required to convert prevalence in primigravida to obtain an estimate for childless women, whereas 110% of the prevalence rate for multigravida could be used to represent prevalence in mothers in the general population. The adjustment factors for high contraception populations are based on results from a single study, representing the experience of 5130 women living in Manicaland, an area of rapidly rising HIV prevalence. A full test on an independent dataset categorized by birth interval has yet to be made. However, the general plausibility of this model can be judged by applying the broad adjustment factors to data from Khutsong district of Gauteng province in South Africa. Ante-natal prevalence in this area in 1999 was 30% in primigravida and 51% in multigravida [32]. About 35% of the women in this community are childless [33]. Applying the correction factors of 0.6 and 1.1 to ante-natal prevalences in the primigravida and multigravida respectively and allowing for their proportionate representation in the female population gives an adjusted overall HIV prevalence estimate of 43%, very close to the overall ante-natal prevalence of 42%. The Carletonville study estimated that community-based HIV prevalence for females 15–44 in 1998 for the same area was 41% [34].

Discussion

We describe a workable method for adjusting ante-natal clinic data to improve estimation of HIV prevalence among women in sub-Saharan Africa. Application of the proposed procedure to data from two East African populations, for which HIV prevalence rates are already known, suggests that it works well. These are populations where the use of contraception is low. Less information is currently available for high contraception-use populations, however, and the appropriateness of this method is less certain. Further assessment of the approach in other areas where community-based surveys have been conducted will have to wait until ante-natal clinic surveillance data are published categorized by parity and birth interval. It is hoped that those responsible for providing advice and support to national sentinel surveillance systems will promote the recording of these extra data items, especially as such information is already collected in the ante-natal clinic setting.

The relative prevalence procedure as presented above does not produce age-specific results. In theory, the method could be made specific for age, and applied separately by age group. In practice, small number problems would make it difficult to obtain estimates of relative prevalence or even the distribution of the population across the different risk categories. We are not proposing that age-based monitoring be abandoned in favour of a parity-based approach, as age-based data are needed for continuity and comparability, and as a guide to incidence at young ages. Prevalence in primigravida is usually very close to prevalence in teenage mothers [26].

The results discussed above suggest that in the cases of Kisesa and Masaka, adjustment for infertility is more important than adjustment for subfertility, the latter accounts for between a quarter and a fifth of the total adjustment. This pattern might vary from one population to another and is likely to depend on the maturity of the epidemic. If subsequent validation checks find that the parity-based adjustment for infertility is much more important than the birth interval-based adjustment for subfertility, it might not be worth recording birth interval data on sero-samples obtained in ante-natal clinics.

Further biases may arise because not all pregnant women avail themselves of the services of ante-natal clinics. This adjustment procedure does not make use of routinely available age data and only makes use of the parity information to draw a distinction between primigravida and multigravida. However, the birth interval information could be used to obtain weighted age and parity distributions of the ante-natal clinic attendees to compare with the corresponding distributions of women in the DHS who gave birth in the last 5 years. Such comparisons would allow the possibility of investigating the presence of selection bias. Whenever community-based HIV prevalence studies are conducted, women could be asked to report not only on their recent fertility, but also on whether or not

they attended an ante-natal clinic for a check up during their last pregnancy. However, the results of such inquiries would be hard to generalize from one community to another, and would not address the question of selection of clinics for participation in the surveillance programme.

The most readily available source of data for obtaining a population breakdown in the childbearing risk categories are DHS surveys, which have been carried out relatively recently in most of the high HIV prevalence countries in Africa for which this adjustment procedure might be relevant. The required tables are not available in the published country reports, but are easy enough to generate from the basic recode files, which are available on the World Wide Web. It might be possible to produce approximations to these distributions from other demographic surveys or even censuses, but since these rarely ask direct questions about sexual activity, it would be necessary to settle for categories such as 'never married' and 'not currently in union' to represent 'never had sex' and 'not currently sexually active'.

The choice of a 5-year reference period for defining infertility could have an impact on the outcome of the adjustment procedure. The longer the reference period, the larger the group we presume to have captured 'directly' by weighting the experience of the pregnant women, and the smaller the corresponding 'infertile' group. In theory this should not matter too much, if we can get reliable estimates of the HIV prevalence for this small group. In the three studies used above, prevalence estimates for the infecund groups look sensible in the light of what is known about HIV and fertility. But in some other application of this method, there might be a strong reason for using the relative prevalence estimates from some other study (because of geographical proximity, for example, or a closer match in terms of contraception use patterns). If the number of sexually active, non-contracepting mothers with birth intervals longer than 5 years is too small to calculate their prevalence level with meaningful confidence, it would be best to redefine the reference period, e.g. as 4 or even 3 years, and tabulate the population distribution from DHS or some other national source in the same way.

The majority of the evidence cited above about HIV prevalence differences between ante-natal clinic sources and population-based studies comes from East and Central African countries which have very low rates of contraception use. Relatively little information on prevalence differentials is yet available from countries such as Kenya, Botswana or South Africa where family planning is more widespread. The multicentre study on factors determining the differential spread of HIV in African cities [35] found no significant difference in ante-natal and community HIV prevalence in Kisumu, Kenya, even though contraception use in Kisumu was not particularly high – only 18% compared with 50% in urban areas of Kenya as a whole, recorded in the 1993 DHS survey [7]. But this does tie in with our limited tests of this method, which suggest that for high contraception use populations it might be acceptable to use unadjusted ante-natal HIV prevalence to represent prevalence in the general female population aged 15–44 years, in contrast to low contraception use populations, where adjustments have a significant impact.

Until more results are available from other study sites we can only speculate on the probable reasons for the differences in the relative prevalence patterns between high and low contraception use populations indicated in Table 2. Part of the explanation may be that in low contraception use populations women in regular partnerships are found mainly in the 'currently fertile' categories, whereas in high use populations those in regular partnerships are just as likely to be contraception users as to be 'currently fertile'. In low use populations, infertile women are relatively easy to identify, but in high use populations some infertile women will be unaware of their infertility and may practice contraception, and contraception users will encompass a much wider socio-economic spectrum. In Manicaland almost 48% of sexually active women are contraception users, whereas in the combined Kisesa and Masaka 'standard' population only 6% use modern contraceptive methods.

Several other factors may merit further investigation, such as the build up of sexual activity with age, the prevalence of other sexually transmitted infections, and the recency of the epidemic. Sexual activity tends to start earlier in less developed countries – the median age at first sex is 16 years in Masaka and Kisesa, whereas in Manicaland it is 18 years. As a result only one-third of childless women aged 15–49 years in Kisesa have never had sex, whereas in Manicaland nearly 70% of childless women are virgins. The proportions of women who appear to be affected by primary and secondary sterility are much higher in Kisesa and Masaka, suggesting higher levels of untreated sterilizing sexually transmitted infections. Although community prevalence levels in Manicaland have reached over 25.6%, far higher than in Kisesa (7.3%) and Masaka (10.5%), the epidemic in the regions bordering Lake Victoria is thought to be older, having started in the early 1980s, whereas that in Eastern Zimbabwe probably started in the late 1980s or early 1990s [36,37]. For this reason, the full fertility reducing effects of HIV will not have had a chance to manifest themselves among infected women in rural parts of Manicaland, as the proportion of recent infections will be higher for these women than for their counterparts in Kisesa or Masaka.

A crucial test for the new method will be the application of the relative prevalence rates shown in Table 2 to data from ante-natal clinic surveillance, suitably classified by parity and birth interval, in communities with varying levels of contraception use, and in areas where the resulting adjusted estimate could be compared to a community-based estimate from the same population. To discover whether the results from Manicaland can be generalized to other more developed populations, it will be necessary to conduct some community-based sero-prevalence studies in countries with higher levels of contraception use, simultaneously collecting data on fertility and its proximate determinants. If it turns out that the relationship between HIV prevalence in contraception users and prevalence in pregnant women varies widely in populations with moderate to high levels of contraception use, it may become necessary to add family planning clinics [38] to the sentinel surveillance centres for such populations, in order to obtain direct evidence of prevalence levels among contraception users.

Surveillance data might also be enriched in other ways – enquiring about mothers' usual residence, or length of residence in the clinic catchment area might give some indication of selection effects due to population mobility. An extension of surveillance to cover births in smaller rural clinics, or even home births attended by traditional birth attendants might present serious logistical problems, but is probably the only way to get over the problems of social selection for attendance in the larger clinics.

Acknowledgements

The authors thank A. Buve, J. Glynn, G. Santow, I. Timaeus and J. Todd for extensive and helpful comments on earlier drafts.

Sponsorship: Supported in part by the UK Department for International Development (Knowledge Programme on reproductive health) (B.Z.), and the USAID-sponsored MEASURE evaluation project, USAID Co-operative agreement #HRN-A-00-97-00018-00.

References

1. United States Bureau of the Census. *HIV/AIDS Surveillance Data Base.* (See also the companion web site, http://www.census.gov:80/ipc/www/hivaidsn.html.) Washington, DC: U.S. Bureau of the Census; 1999.
2. Statistics Department and Macro Internationa [Uganda]. *Uganda Demographic and Health Survey, 1995.* Calverton, Maryland: Statistics Department and Macro International; 1996.
3. Central Statistics Office and Macro International [Botswana]. *Botswana Family Health Survey II 1988.* Calverton, Maryland: Central Statistical Office and Macro International; 1989.
4. Bureau of Statistics and Macro International [Tanzania]. *Tanzania Demographic and Health Survey 1996.* Calverton, Maryland: Bureau of Statistics and Macro International; 1997.
5. Central Statistical Office and Macro International [Zimbabwe]. *Zimbabwe Demographic and Health Survey 1994.* Calverton, Maryland: Central Statistical Office and Macro International; 1995.
6. Central Statistical Office and Ministry of Health and Macro International [Zambia]. *Zambia Demographic and Health Survey 1993.* Calverton, Maryland: Central Statistical Office and Macro International; 1994.
7. National Council for Population and Development, Central Bureau of Statistics and Macro International [Kenya]. *Kenya Demographic and Health Survey 1993.* Calverton, Maryland: NCPD, CBS and MI; 1994.
8. Carpenter L, Nakiyingi J, Ruberantwari A, Malamba S, Kamali A, Whitworth J. **Estimates of the impact of HIV infection on fertility in a rural Ugandan population cohort.** *Health Trans Rev* 1997, **7** (suppl 2):113–126.
9. Gray R, Wawer M, Serwadda D, *et al.* **Population-based study of fertility in women with HIV-1 infection in Uganda.** *Lancet* 1998, **351**:98–103.
10. Boerma JT, Urassa M. **Associations between female infertility, HIV and sexual behavior in a rural area in Tanzania.** In *Women and Infertility in Sub-Saharan Africa: A Multidisciplinary Perspective.* Edited by Boerma JT, Mgalla Z. Amsterdam: Royal Tropical Institute; 2000: 176–187.
11. Temmerman M, Plummer F, Mirza N, *et al.* **Infection with HIV as a risk factor for adverse obstetrical outcome.** *AIDS* 1990, **4**:1087–1093.
12. Miotti P, Dallabetta G, Ndovi E, Liomba G, Saah A, Chiphangwi J. **HIV-1 and pregnant women: associated factors, prevalence, estimate of incidence and role in fetal wastage in central Africa.** *AIDS* 1990 **4**:733–736.
13. De Cock K, Zadi F, Adjorlolo G, *et al.* **Retrospective study of maternal HIV-1 and HIV-2 infections and child survival in Abidjan, Côte d'Ivoire.** *BMJ* 1994, **308**:441–443.
14. Brocklehurst P, French R. **The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis.** *Br J Obstet Gynaecol* 1988, **105**:836–848.
15. Ross A, Morgan D, Lubega R, Carpenter LM, Mayanja B, Whitworth J. **Reduced fertility associated with HIV: the contribution of pre-existing subfertility.** *AIDS* 1999, **13**:2133–2141.
16. Zaba B, Collumbien M. **HIV and fertility: modeling the effects of changes in union dynamics.** *Proceedings of the IUSSP General Conference, Beijing 1997.* Liege: Ordina Editions:583–610.
17. Fylkenses K, Ndholovu Z, Kasumba K, Musonda RM, Sichone M. **Studying dynamics of the HIV epidemic: population-based data compared with sentinel surveillance data.** *AIDS* 1998, **12**: 1227–1234.
18. Kigadye RM, Klokke A, Nicoll A, *et al.* **Sentinel surveillance for HIV-1 among pregnant women in a developing country: 3 years' experience and comparison with a population serosurvey.** *AIDS* 1993, **7**:849–855.
19. Zaba B, Gregson S. **Measuring the impact of HIV on fertility in Africa.** *AIDS* 1998, **12**(suppl 1):S41–S50.
20. Lessner L. **Projection of AIDS incidence in women in New York state.** *Am J Publ Health* 1991, **81** (suppl):S30–S34.
21. Boisson E, Nicoll A, Zaba B, Rodrigues LC. **Interpreting HIV seroprevalence data from pregnant women.** *J Acquir Immune Def Syndr Hum Retrovirol* 1996, **13**:434–439.
22. Nicoll A, Stephenson J, Griffioen A, Boisson E, **The relationship of HIV prevalence in pregnant women to that in women of reproductive age: a validated method for adjustment.** *AIDS* 1998, **12**:1861–1867.
23. Karon MJ. **Methods for estimating HIV prevalence in the United States.** Technical report no. D821. Atlanta: Centers for Disease Control and Prevention; 1997.
24. Lee L. **Fertility reduction and duration of HIV infection: Findings from the United States.** In *HIV, STI and Infertility: Past Trends and Current Monitoring Problems. Conference abstracts, USAID, MEASURE Evaluation and UNAIDS.* Arlingtom, Virginia: Measure Evaluation;1998.
25. Garnett G, Gregson S. **Monitoring the course of the HIV epidemic: the influence of patterns of fertility on HIV prevalence estimates.** *Mathematical Population Studies* 2000, **8**: (in press).

26. Zaba B, Boerma JT, White R. **Monitoring the AIDS epidemic using HIV prevalence data among young women attending antenatal clinics: prospects and problems.** *AIDS* 2000, **14**:1633–1645.
27. Blacker JGC, Afzal M, Jalil A. **The estimation of fertility from distributions of births by order: application to the Pakistan demographic survey.** *In Proceedings of the IUSSP International Population Conference, vol. 2.* Liege: Ordina Editions; 1989:103–111.
28. Glynn JR, Zaba B, Carael M. **HIV and Fertility: report on proceedings of a UNAIDS workshop.** Presented at *The Joint United Nations Programme on HIV/AIDS.* London, May 1998.Geneva: UNAIDS.
29. Larsen U. **Sterility in sub-Saharan Africa.** *Population Studies* 1994, **48**:459–474.
30. Boerma JT, Urassa M, Senkoro K, Klokke A, Ng'weshemi JZL. **Spread of HIV infection in a rural area in Tanzania.** *AIDS* 1999, **13**:1233–1240.
31. Gregson S, Zhuwau T, Anderson RM, Chandiwana SK. **Is there evidence for behaviour change in response to AIDS in rural Zimbabwe?** *Soc Sci Med* 1998, **46**:321–330.
32. Makubalo L, Levin J, Zaba B, Mulumba R, Bikitsha N, Mahlasela L. **Recent trends in HIV prevalence in South Africa.** *XIII International Conference on AIDS.* Durban, July 2000 [abstract LbPp121].
33. Department of Health and Macro International [South Africa]. *South Africa Demographic and Health Survey 1998: preliminary report.* Pretoria: Department of Health and Macro International; 2000.
34. Gilgen D, Campbell C, Williams B, Taljaard D, MacPhail C. **The Natural History of HIV/AIDS in South Africa – a biomedical and social survey in Carletonville.** Johannesburg: CSIR; 2000.
35. Kahindo M, Nyang J, Hawken MP, *et al*. **Comparison of HIV-1 seroprevalence in antenatal clinics and the general population in Kenya.** *XI International Conference on AIDS and STIs in Africa.* Lusaka, Zambia, September 1999 [abstract 14ET3-6].
36. Gregson S, Zhuwau T, Anderson RM, Chandiwana SK. **HIV and fertility change in rural Zimbabwe.** *Health Trans Rev* 1997, **7**:(suppl 2):89–112.
37. Gregson S, Zhuwau T, Anderson RM, Chandiwana SK **The early socio-demographic impact of the HIV epidemic in Rural Zimbabwe.** Oxford: Blair Research Institute and Oxford University; 1996.
38. Kapiga SH, Lyamuya EF, Lwihula GK, Hunter DJ. **The incidence of HIV infection among women using family planning methods in Dar es Salaam, Tanzania.** *AIDS* 1998, **12**:75–84.