Uncertainty in estimates of HIV/AIDS: the estimation and application of plausibility bounds

N C Grassly, M Morgan, N Walker, G Garnett, K A Stanecki, J Stover, T Brown, P D Ghys

Objectives: To establish the accuracy of the country specific estimates of HIV prevalence, incidence, and AIDS mortality published every 2 years by UNAIDS and WHO.

Methods: We review sources of error in the data used to generate national HIV/AIDS and where possible estimate their statistical properties. We use numerical and approximate analytic methods to estimate the combined impact of these errors on HIV/AIDS estimates. Heuristic rules are then derived to produce plausible bounds about these estimates for countries with different types of epidemic and different qualities of surveillance system.

Results: Although 95% confidence intervals (CIs) can be estimated for some sources of error, the sizes of other sources of error must be based on expert judgment. We therefore produce plausible bounds about HIV/AIDS estimates rather than statistical CIs. The magnitude of these bounds depends on the stage of the epidemic and the quality and coverage of the sentinel HIV surveillance system. The bounds for adult estimates are narrower than those for children, and those for prevalence are narrower than those for new infections.

Conclusions: This paper presents a first attempt at a rigorous description of the errors associated with estimation of global statistics of an infectious disease. The proposed methods work well in countries with generalised epidemics (>1% adult HIV prevalence) where the quality of surveillance is good. Although methods have also been derived for countries with low level or concentrated epidemics, more data on the biases in the estimation process are required.

See end of article for authors' affiliations
....

Correspondence to:
Dr P D Ghys, UNAIDS, 20 Avenue Appia, CH-1211 Geneva 27, Geneva, Switzerland;
ghysp@unaids.org
....

U NAIDS and the World Health Organization (WHO) have produced country specific estimates of HIV/AIDS and its impact every 2 years since 1998.[1-4] These estimates have been useful in developing awareness of the global extent of HIV/AIDS and have been a motivation for increased focus and spending in combating the HIV/AIDS epidemic. At the same time, the estimates of HIV prevalence and AIDS deaths have come under increasing scrutiny, and have been criticised as too high, too low, or simply the result of political expediency. While some of the criticism is justified and some misguided, much of the problem has stemmed from a misunderstanding of the estimates: how they are made, how precise they are, and why they change over time. Other papers in this supplement focus on the methods and assumptions used to make estimates of HIV/AIDS. This paper focuses on describing the uncertainty around the estimates of HIV/AIDS and how bounds about the UNAIDS/WHO estimates for the end of 2003 were developed.

In this paper we briefly review the methods and assumptions that UNAIDS and WHO have used to make estimates of HIV/AIDS and the evidence for the error or variance around each step in the estimation process. For each step we estimate the associated statistical error where feasible and then introduce judgements of plausibility based on the generalisability of these statistical errors. We derive methods to combine these various sources of error to produce the plausibility bounds around the estimates of HIV/AIDS. Different methods are used for countries with generalised epidemics where adult HIV prevalence exceeds 1% and for low level and concentrated epidemics where prevalence is lower but can exceed 5% among at-risk groups in the case of the latter. The paper concludes with a discussion of how new data sources can be used to further refine the estimates and future directions in this work.

METHODS AND RESULTS

Two different approaches are used to make estimates of HIV/AIDS depending on the type of epidemic, generalised or low level/concentrated. These approaches are described in detail in other papers in this volume.[5-7] We describe error associated with each step and assumption separately for the two approaches.

Adult estimates for generalised epidemics

In making a national estimate of adult (15–49 years old) HIV prevalence, new infections, and AIDS mortality there are four primary sources of uncertainty. These are uncertainty in the estimate of:

- adult (15–49 years old) prevalence from a sample of pregnant women at an antenatal clinic (ANC)
- survival of infected adults
- the epidemic curve used to estimate new infections and mortality from prevalence data
- national prevalence from the typically non-random sample of ANCs (fig 1).

Estimates of HIV prevalence, new infections, and AIDS mortality in children (due to mother to child transmission; MTCT) have additional assumptions and sources of error, which are reviewed in later sections.

1) Estimating adult prevalence from pregnant women at ANCs

Estimates of adult (15–49 year old men and women) HIV prevalence from community studies were reviewed and

....

Abbreviations: ANC, antenatal clinic; ART, antiretroviral therapy; MTCT, mother to child transmission
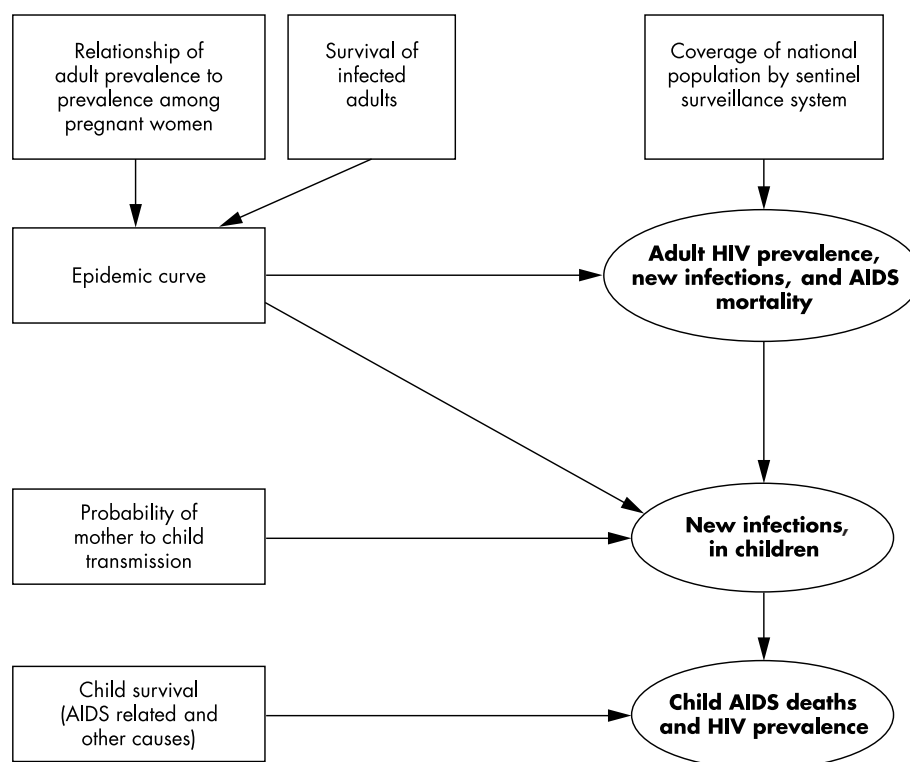
Figure 1 Sources of uncertainty in the steps used to make estimates of HIV/AIDS in countries with generalised epidemics. Boxes represent sources of uncertainty while ellipses represent estimates produced by UNAIDS/WHO. The arrows indicate which estimates are affected by different sources of uncertainty.

compared with prevalence estimates from ANCs in the same locality at the same time (maximum 1 year difference) (table 1). Only three of the 15 comparisons in eight countries show any significant difference between ANC and community prevalence. Prevalence from the ANC sample overestimated community prevalence in studies in 1997 in Rwanda and Kenya, and underestimated community prevalence in a 1999 study in Zambia. It is possible from these comparisons to explore the statistical attributes of the relationship between ANC and adult population prevalence.

If pregnant women at ANCs represent a random sample of the local population with respect to HIV prevalence, the number of HIV positive samples X should be binomially distributed with mean ρn, where ρ is the population prevalence and n the sample size. This hypothesis can be tested by comparing the fit of the binomial distribution to the samples described in table 1 with a distribution that allows additional variance and the mean of the ANC sample to differ systematically from the community prevalence. An appropriate distribution is the beta binomial, typically used in Bayesian analysis. In this case the observed number of HIV cases in the sample is drawn from a binomial distribution with a mean θ, which itself is drawn from a beta distribution with a mean that is a function of the community prevalence. In the simplest case the mean of the beta binomial $\mu = \gamma\rho$, where γ is the prevalence of the ANC sample relative to

Table 1 Antenatal clinic (ANC) and community adult HIV prevalence

	ANC		Population		ANC:		
Location and year	prevalence	n	prevalence	n	community ratio	CIs	Reference
Fort Portal, Uganda, 95	18.4%	477	22.9%	875	0.80	0.64–1.01	Kilian et al[8]
Chelston, Zambia, 95–6*	26.1%	532	25.7%	1909	1.02	0.86–1.20	Fylkesnes et al[9]
Chelston, Zambia, 99*	25.9%	776	23.0%	1768	1.12	0.97–1.30	Fylkesnes et al[10]
Kapiri Mposhi, Zambia, 95–6*	12.6%	422	16.7%	760	0.76	0.56–1.02	Fylkesnes et al[9]
Kapiri Mposhi, Zambia, 99*	8.3%	300	16.7%	724	0.50	0.33–0.75	Fylkesnes et al[10]
Yaounde, Cameroon, 97–8	5.5%	1532	6.1%	1913	0.90	0.69–1.18	Glynn et al[11]
Kisumu, Kenya, 97–8	30.6%	1480	25.9%	1515	1.18	1.05–1.33	Glynn et al[11]
Ndola, Zambia, 97–8	27.3%	1021	28.4%	1534	0.96	0.85–1.09	Glynn et al[11]
urban Mwanza, Tanzania, 90–1	11.7%	1820	11.8%	1184	0.91	0.76–1.08	Kigadye et al[12]
Manicaland, Zimbabwe, 98–00†	21.1%	1215	22.5%	9119	0.94	0.84–1.05	Gregson et al[13]
Manicaland, Zimbabwe, 01–03†	19.2%	1237	19.9%	7153	0.96	0.85–1.09	Personal communication‡
Kagera, Tanzania, 93	17.3%	2816	18.1%	653	0.96	0.80–1.15	Kwesigabo et al[14]
Kagera, Tanzania, 96	13.0%	2893	12.9%	1276	1.01	0.85–1.20	Kwesigabo et al[14]
Rwanda, 97†	15.1%	1368	12.7%	3030	1.19	1.02–1.39	Ministry of Health[15]
Karonga, Malawi, 98–01	10.4%	2998	12.6%	833	0.83	0.67–1.02	Crampin et al[16]

*Community survey 15–39 year olds. †Community survey 15–44 year olds.‡Simon Gregson, personal communication February 2004.

community prevalence. Thus, for instance, if prevalence at ANC systematically underestimates prevalence in the community then $\gamma < 1$. We re-parameterise the beta distribution from $beta\ (a,b)$ to $beta\ (\mu, M)$, where $\mu = a/(a+b)$ the population mean and $M = a+b$, a measure of extra variance (increasing M decreases the extra variance). We therefore have a two stage compound sampling model:

$$\theta \sim beta\ (\mu, M)$$
$$X|\theta \sim Bin(n, \theta)$$

The beta binomial has the nice property that the expected value of X/n is:

$$E(X/n) = E[E(X/n|\theta)] = E[\theta] = \mu$$

and the variance:

$$V(X/n) = [1+(n-1)/(M+1)]\ \mu(1-\mu)/n$$

where $(n-1)/(M+1)$ represents the extra binomial variance.

The fit of the binomial model to the data in table 1 can be compared to the fit of the beta binomial using the likelihood ratio statistic. We test three hypotheses, namely that prevalence among pregnant women at ANC represents: H_0 – a random sample of prevalence in the community (binomial sample); H_1 – a random sample of prevalence in the community but with additional variance (beta binomial, $M << \infty$, $\gamma = 1$); and H_2 – a biased sample of prevalence in the community (beta binomial, $M << \infty$, $\gamma \neq 1$). Because these hypotheses are nested within one another, twice the likelihood ratio is approximated by the χ^2 distribution with the number of degrees of freedom equal to the difference in the number of parameters.

Compared to hypothesis H_1 the null hypothesis H_0 of a random binomial sample is rejected ($p < 0.001$). However, although the ANC sample slightly underestimates community prevalence ($\hat{\gamma} = 0.94$), the difference is not significant ($p = 0.09$). This provides strong empirical support for the assumption made by UNAIDS that prevalence among pregnant women represents a good estimate of adult (men and women combined) prevalence in the local community. However, the variance of the estimate is greater than that expected from a simple random sample ($\hat{M} = 221$). For a sample size $n = 300$, the variance of the sample would be an estimated 2.3 times that expected from a random binomial sample.

2) Adult survival after infection with HIV

Median adult survival without access to antiretroviral treatment or intensive case management with treatment for opportunistic infections was estimated to be 9 years, with a standard deviation of 1 year. This is based on published reviews of cohort studies before the availability of highly active antiretroviral therapy.[17][18] The distribution of survival is described by a Weibull distribution with a fixed shape parameter previously estimated.[17] The estimate of adult survival is not only important in the estimation of adult AIDS mortality but also determines the shape of the prevalence and incidence curves. Uncertainty in survival was therefore included in estimates of the range of possible epidemic curves.

3) The epidemic curve

HIV prevalence estimates at a given point in time provide information about likely values of prevalence in the following year. Time series of HIV prevalence at ANCs therefore provide more information about HIV prevalence at a given point in time than do estimates based on a single year of reporting. UNAIDS and WHO have developed various approaches to fitting an epidemic trajectory to sampled prevalence data. Currently a simple model of HIV transmission among a susceptible fraction of the adult population is used to fit prevalence data.[17] This model is capable of providing a good fit to a range of epidemics with just four parameters

including the start date of the epidemic. Estimates of new infections and AIDS mortality are necessarily derived from this model because independent data to estimate these are exceedingly rare.

Data from a given country may be consistent with a range of epidemic curves. Because we have an estimate of the distribution of uncertainty about ANC estimates and survival after infection, it is possible to calculate the likelihood of a given epidemic curve. The likelihood function depends strongly on the assumptions made about whether the distribution of prevalence estimates from ANC surveillance sites in a country is independent. If one surveillance site in a country tends to underestimate community prevalence, other sites may be more likely to do so. This covariance will increase the uncertainty in the estimated epidemic curve over that seen if sites were independent. Covariance of estimates from sites can be modelled by specifying correlation in the mean of the binomial sample (θ) given by the beta prior.

Within the likelihood framework, any curve that is not significantly different from the maximum likelihood is considered to be within the "set of admissible hypotheses".[19] This set of hypotheses is often considered a close approximation of the confidence interval (CI), although they do not represent a CI in the strictest sense, where the type I error rate should equal the significance level, α. However, there is no simple analytic approach to relate uncertainty in the UNAIDS model parameters to uncertainty in prevalence, new infections, and mortality. It is also difficult to numerically find "admissible" curves that give maximum and minimum estimates of prevalence at a given time. We therefore adopt a parametric bootstrap approach to explore the range of curves consistent with the data. Observed prevalence data were re-sampled using the beta binomial distribution with $M = 221$ (as estimated above), $n = 300$, and mean equal to the observed prevalence. A median survival time is randomly chosen from the normal distribution with a mean of 9 years and standard deviation of 1 year. The UNAIDS model epidemic curve is then fit using a goodness of fit measure (in this case least squares to match the implementation in the Estimation and Projection Package (EPP) software used by UNAIDS). This is then repeated 1000 times and the distribution of adult prevalence, new infections, and AIDS mortality recorded. Three models of covariance of sites were explored: each ANC prevalence estimate is independent (m_0), sites are independent but estimates for a given site over time are correlated (m_1), sites in a given country or region are correlated over time and with each other (m_2). To implement the parametric bootstrapping procedure random deviates for θ were derived from the beta distribution by solving $U[0..1] = beta\ cdf\ (\mu, M)$, where $U[0..1]$ is a random number uniformly distributed between 0 and 1, and $beta\ cdf\ (\mu, M)$ is the cumulative probability density function of the beta distribution. We use the same uniform random deviate for different sites (each with a different associated community mean μ) to allow for correlation in θ. In reality the covariance of sites is likely to lie somewhere between m_1 and m_2.

Fig 2 shows an example application of this approach to prevalence data from urban Zambia for covariance model m_2. As expected, uncertainty in the estimates of new infections and AIDS mortality is greater than for adult prevalence, reflecting the additional uncertainty in adult survival and the shape of the epidemic curve.

It is not practical to apply this approach to all sets of prevalence data that are used to produce HIV/AIDS estimates in countries with generalised epidemics. We therefore adopt a heuristic approach to derive some general rules from applying the bootstrap approach to a representative set of eight countries. We classified these countries by quality of surveillance: good (Uganda, Zambia, and Ghana); average
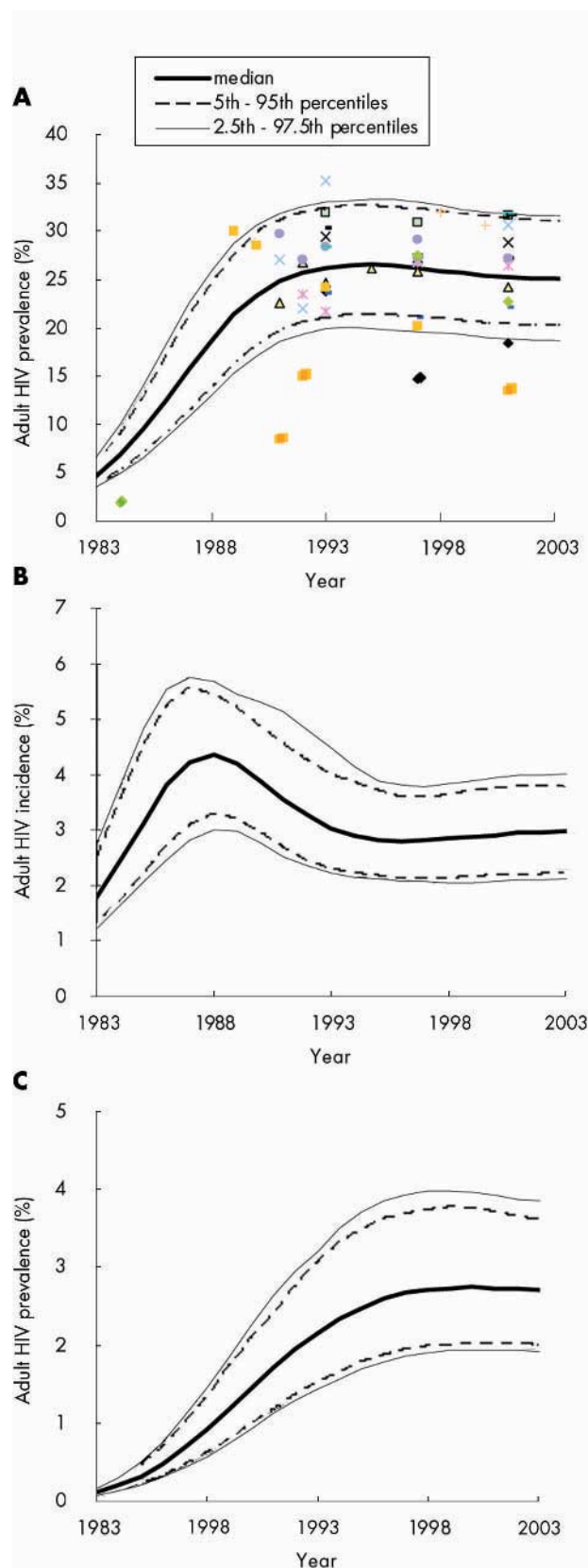
Figure 2 The range about the median adult. (A) HIV prevalence, (B) HIV incidence, and (C) AIDS mortality estimates over time for urban Zambia for 1000 parametric bootstrap replicates assuming error in the mean of the sample from each site is correlated (covariance model m₂). Zambia has a surveillance system classified as good.

(Democratic Republic of the Congo and Chad); and poor (Eritrea, Angola, and Djibouti).[20] We then applied the bootstrap approach to urban and rural data for each country.

Countries with poor surveillance and rural areas of countries with moderate surveillance have limited prevalence data, reflected in a vast range of epidemic curves that are statistically consistent with the data. In reality, a large fraction of these curves would be discarded by the epidemiologist making the estimate as being unrealistic. This judgement would usually be based on analogy with similar countries or consideration of what is possible epidemiologically. These criteria are not, therefore, provided by statistical analysis of the prevalence data but extrinsic considerations. This necessitates the use of plausible rather statistical bounds on estimates. Countries with limited data also tend to have poor coverage of more rural areas, further making plausibility the only sensible criterion. Plausible bounds for these countries are discussed further in the next section.

Countries with good surveillance and urban areas of countries with average surveillance have a narrower range of prevalence curves (see Zambia, fig 2). The 95% CIs for adult prevalence, new infections, and mortality are asymmetric (larger upper bound) and proportionally bigger for smaller estimates. This suggests the use of the logit transform ($\ln(x/(1-x))$) where x is the estimate as a proportion of the adult population. This transformation results in symmetric, approximately normally distributed bounds about the estimates of prevalence, new infections, and mortality. For a given model of covariance of error across sites, the coefficients of variation (c.v. = standard deviation/mean) for each of the three transformed estimates are broadly similar for the countries with good surveillance data irrespective of prevalence level when calculated at the same point in time with respect to the latest available prevalence data. Thus, for example, estimates of adult HIV prevalence in 2003, 1 year after the latest available data, have broadly similar c.v. for Uganda, Zambia, and Ghana. Estimates for 2001 have smaller c.v. because prevalence data are available for the year of the estimate as well as the following year. The only exception to this pattern is adult AIDS mortality, which has less uncertainty (smaller c.v.) in countries where prevalence is declining or has been steady for some time. This is because AIDS mortality in a given year is largely dependent on prevalence levels 5–10 years earlier, and in these countries earlier prevalence estimates are better specified.

The appropriate model of covariance in the error in different ANC sites reporting at different times is unknown. It is likely to lie between models m₁ and m₂, as previously discussed, and in developing a heuristic rule we simply take the average of the c.v. for these two models. Further, we assume that the c.v. for a national estimate is the average of the c.v. for the urban and rural epidemic curves weighted by the population numbers in these areas. The heuristic, therefore, consists of a set of c.v. on the logit scale that are applied to country estimates of adult HIV prevalence, women living with HIV/AIDS, new infections, and AIDS mortality based on the quality of the surveillance system and the stage of the epidemic (rising, steady, and falling).

4) National estimates from a non-random sample

Although the method described above accounts for the quantity of data over time, it does not capture the representativeness of the sites with regard to the national population. Historically, the sentinel sites in countries in sub-Saharan Africa have been placed in urban and peri-urban areas. This ensures a larger number of pregnant women to test during the surveillance period and also makes logistics of testing and collation of surveillance data easier. Unfortunately, it also means that most surveillance systems

do not provide much data on prevalence among pregnant women in rural areas.

This bias in the surveillance data cannot be accurately measured. In previous estimates for countries in sub-Saharan Africa, we have generally reduced rural prevalence by 20% to adjust for the bias. This adjustment can be improved upon by local review of clinic location and catchment population along with detailed demographic data on the geographic distribution of the entire population. This is only possible through laborious local consultation as carried out recently in Kenya and Malawi.[21] [22] The amount of coverage of rural pregnant women provided by the sentinel system varies by country. In an effort to capture this uncertainty and possible bias we have chosen to expand the uncertainty associated with the estimates based on the quality of coverage provided by the data.

In the absence of detailed local consultation, the plausibility bounds for countries where rural coverage is poor have been expanded by an additional 10% of their value. This includes most countries where the surveillance system is classified as poor or average.

Child estimates for generalised epidemics
The number of infants born with HIV and the number of children aged <15 years olds living with HIV/AIDS and dying from AIDS are estimated for each country with a generalised epidemic. These estimates are based on prevalence among pregnant women (equivalent to the adult prevalence estimates from ANCs), the probability of MTCT, and life tables for the survival of HIV positive children (fig 1). There is greater uncertainty in these estimates than for adult prevalence alone. Methods to estimate the additional uncertainty in child estimates for generalised epidemics and to produce plausible bounds about estimates are described below. Example plausibility bounds derived using these heuristics for countries with different quality surveillance systems are shown in table 2.

Incident child HIV infections
The estimate of the fraction of live births that are HIV positive is simply the product of ANC prevalence and the probability of MTCT. The latter is assumed to be 0.32 with a standard deviation of 0.05 based on a recent review.[17] We use the delta method to combine uncertainty in these two estimates to obtain overall uncertainty in the fraction of births that are HIV positive.[23] Numbers of incident child infections are simply obtained by multiplying by the number of live births in a country and we ignore any uncertainty in this quantity.

Child AIDS deaths
The number of deaths among children in a country for a given year is derived from the convolution of the trend in HIV positive births and child mortality due to AIDS and other causes. Estimates of uncertainty in some of these parameters may be derived, while for others the uncertainty is largely unknown. We have chosen an approach that allows us to use available data to include some of the uncertainty in the estimates when producing the plausibility bounds, while other sources of uncertainty have been ignored. As such the plausibility bounds presented for child AIDS deaths are probably too narrow.

Firstly, consider the uncertainty in HIV related mortality among infants and children. The UNAIDS Reference Group has taken child survival from a five parameter double Weibull curve fit.[17] The parameters in the curve were estimated by a non-linear least squares fit to published data. To estimate the uncertainty in HIV survival, those same data were refit to the double Weibull, again using least squares, and the parameter covariance matrix was estimated. Then a sample of 1000 sets of double Weibull parameters was drawn from a multivariate normal distribution based on that covariance matrix.

The distribution of HIV survival curves was then combined with a distribution of curves for survival from mortality due to other causes using Monte Carlo simulation to estimate the c.v. for the number of deaths due to AIDS in 2003 conditional on three patterns of incidence of HIV infections among infants over time (rising, steady, and falling). Mortality due to causes other than AIDS in the first year of life was assumed to be normally distributed with a mean of 10% and standard error of 5%. In subsequent years the hazard was assumed to be half that of the randomly selected hazard used in the initial year of life. Using this approach the c.v. was found to vary somewhat according to the epidemic pattern. For a rising epidemic the c.v. was ~0.065, for a steady epidemic it was ~0.08, and for a declining epidemic it was ~0.09. This is because for a declining epidemic proportionally more AIDS deaths occur at older ages when estimates of survival are more uncertain.

To derive heuristic rules to generate plausible bounds around estimates of child deaths in each country with a generalised epidemic, uncertainty in child AIDS deaths for a given incidence of child HIV infections must be combined with uncertainty in incidence. We make the simplifying

Table 2 Example application of heuristic rules to obtain plausibility bounds for estimates of HIV/AIDS in a country with a stable, generalised HIV epidemic

Best estimate	Plausibility bounds		Relative value	
			Minus	Plus
Good surveillance system, high HIV prevalence				
16.5%	(13.5%–20.1%)	Adult prevalence	18%	22%
69 400	(50 000–97 000)	Adult deaths	28%	40%
95 000	(60 000–150 000)	New adult infections	37%	58%
1.8%	(1.2%–2.7%)	Child prevalence	35%	53%
20 200	(13 600–30 000)	Child AIDS deaths	33%	49%
24 000	(16 600–34 500)	New child infections	31%	43%
Average surveillance system, low HIV prevalence				
4.2%	(1.7%–9.9%)	Adult prevalence	59%	136%
78 400	(40 000–153 000)	Adult deaths	49%	95%
123 600	(33 000–465 000)	New adult infections	73%	276%
0.5%	(0.2%–1.3%)	Child prevalence	62%	160%
26,200	(10 200–67 300)	Child AIDS deaths	61%	157%
34 300	(13 400–86 300)	New child infections	61%	152%

assumption that the number of deaths due to AIDS is a linear function of the number of incident HIV infections. We can therefore use the delta method to estimate uncertainty in the former from the simulated c.v. for the number of deaths given a fixed pattern of incidence and the cv for incidence estimated in the preceding section. Because we are working with numbers rather than proportions, variables are transformed on the log scale.

Child prevalence

The plausibility bounds for the number of children living with HIV are derived in a manner almost identical to that for child deaths. The only difference is that, instead of estimating the c.v. for deaths conditional on the number and pattern of the epidemic, the c.v. for children living with HIV was calculated in the simulation. For a rising epidemic the c.v. for child prevalence was found to be 0.10, for a steady epidemic it was 0.11, and for a declining epidemic it was 0.12.
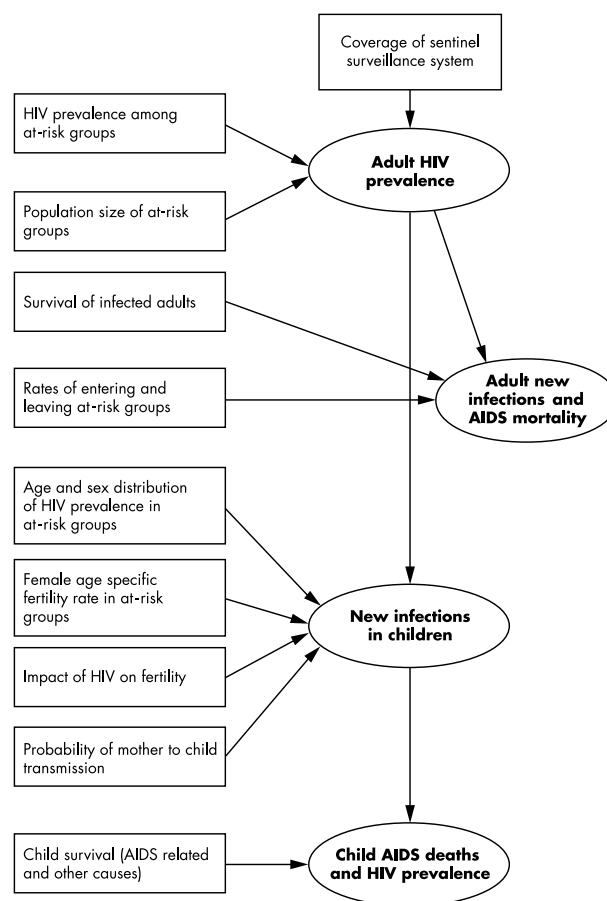
Adult estimates for low level and concentrated epidemics

Estimates of national adult HIV prevalence for countries with low level or concentrated epidemics are based on the sum of the products of the estimated population size and prevalence for all groups considered at higher risk of HIV infection, and for the population at low risk. Estimates of adult new infections and AIDS mortality are subsequently derived by fitting an epidemic curve to a series of national prevalence estimates (fig 3). Those groups at higher risk of infection differ between countries, where epidemics may be concentrated among drug users, sex workers and their clients, men who have sex with men, or a combination of these. The diversity of low level and concentrated epidemics means it is difficult to estimate plausibility bounds that are comparable across countries.

Adult HIV prevalence

Workbooks to assist with the estimation process for low level and concentrated epidemics are described in an earlier paper in this series.[5] These workbooks require low and high estimates of both prevalence and population size for at-risk groups in addition to the best estimate. These low and high estimates are typically based on expert judgement of plausibility, because surveillance data for these quantities are rarely amenable to statistical analysis. The advantage to this approach to capturing error is that it allows local analysts to set the plausibility bounds based on their knowledge of the strengths and weaknesses of their estimates. The disadvantage is that there is likely to be wide variation in how people decide on plausible bounds given their datasets.

   Although statistical CIs can be calculated for estimates of population size for at-risk groups when standard techniques such as capture–recapture or multiplier methods are used, these techniques are rarely employed. Furthermore, where they have been employed, they are typically restricted geographically and therefore judgement must be made as to their appropriateness for application elsewhere in a country to produce a national estimate. Similarly, coverage of HIV sentinel surveillance sites among at-risk groups is typically poor, and judgements about the accuracy of national prevalence estimates must be made. Therefore, for bounds on estimates of adult prevalence and number of adults and women living with HIV/AIDS, we have simply used the median of the bounds provided by countries with low level epidemics (−70% and +100% of the best estimate) and concentrated epidemics (−50%, +60%). These plausibility bounds capture the perceived uncertainty in both the estimates of the sizes of the populations used in making

Figure 3 Sources of uncertainty in the steps used to make estimates of HIV/AIDS in countries with low level and concentrated epidemics. Boxes represent sources of uncertainty while ellipses represent estimates produced by UNAIDS/WHO. The arrows indicate which estimates are affected by different sources of uncertainty.

the overall estimate as well as the uncertainty in the estimates of prevalence in each group.

Adult new infections and AIDS mortality

Estimating adult new infections and mortality in countries with low level or concentrated epidemics is more difficult and less certain than for generalised epidemics. The primary difficulty is reliance on an overall epidemic curve composed from national point estimates in multiple years and the concomitant lack of distinct epidemic curves for each of the at-risk groups. Some at-risk groups are likely to have a different background mortality (for example higher rates of death in drug users), and the duration of stay in the at-risk groups may vary. This can have substantial effects on the relationship between observed HIV prevalence and the underlying patterns of incidence and mortality that are inferred from the prevalence using the epidemic curve. Unfortunately country specific data on mortality and changes in risk behaviour are typically absent. So we have simply used the relationship between uncertainty in HIV prevalence and uncertainty in new infections and AIDS mortality estimated for generalised epidemics and applied it to low level and concentrated epidemics to derive minimum bounds about the estimates. This yields median plausibility bounds of −50% to +70% of the best estimate for adult AIDS deaths and −70% to +190% for adult new infections in countries with concentrated epidemics. For countries with low level epidemics the bounds are −70% to +80% for AIDS deaths and −80% to

+240% for new infections. These bounds, while quite large, do not reflect the full uncertainty associated with all the steps in estimating adult new infections and AIDS mortality (fig 3) and are likely to be too narrow.

Child estimates for low level and concentrated epidemics

The majority of countries with low level or concentrated epidemics do not have estimates of the prevalence of HIV among pregnant women. Estimates for children require several additional steps compared with countries with generalised epidemics (fig 3). Estimates of the age and sex distribution of prevalent HIV infections are used to calculate the number of women with HIV by age group. Age specific patterns of fertility with suitable adjustments for the impact of HIV on fertility are then used to estimate the number of births to HIV positive women.[7] Multiplying by the probability of MTCT gives the number of incident child HIV infections.

The fertility of female sex workers or injecting drug users is likely to differ from that of women in the general population. Unfortunately data on fertility in these groups are currently unavailable. In addition, data on the age distribution of HIV infection in these groups are limited. This makes plausible bounds around estimates of child HIV infections difficult to derive. Estimates of child AIDS deaths and HIV prevalence are based on incident child infections and so plausibility bounds on these estimates are similarly difficult to derive. So, we choose not to publish country specific HIV estimates for children in low level and concentrated epidemics.

DISCUSSION AND CONCLUSIONS
Why plausibility bounds

In this work we have chosen to describe the error around HIV/AIDS estimates in terms of plausibility bounds instead of statistical CIs because not all sources of error are amenable to statistical analysis. Wherever possible, formal statistical procedures have been used to quantify uncertainty. However, the resulting statistical CIs must then be combined with estimates of uncertainty based on expert judgement rather than quantitative analysis because of the absence of appropriate data. For example, biases in estimates of rural HIV prevalence due to the restricted locations of sentinel surveillance sites often cannot be quantified because of the lack of population prevalence surveys with more representative coverage that can provide a comparison. Instead, the size of the bias must be estimated based on consideration of the location of the clinics, rural–urban prevalence differences, and observations from other countries where population surveys have taken place. This use of plausibility bounds is similar to the approach taken in other fields (for example information theory, artificial intelligence)[24 25] where information of varying quality has to be combined. It also has links to work in public health where decisions have to be based on expert judgement in addition to formal statistical inference.[26 27]

Use of other data sources

The HIV/AIDS estimates and their uncertainty explored in this paper are based on the analysis of sentinel surveillance data. Alternative data sources may sometimes be available to help refine the estimates and reduce the plausibility bounds.

Countries with concentrated epidemics often have additional sources of data that can directly inform and constrain estimates, narrowing the bounds to be placed on them. For example, in the Russian Federation all pregnant women are tested for HIV, and so the estimates of the number of children infected via MTCT can be accurate. Likewise, in countries such as Brazil and Argentina that have extensive voluntary

counselling and testing programmes, case reports may also substantially reduce the bounds on the estimates.

In countries with generalised epidemics, estimates of deaths derived from vital registration systems or census rounds can be used to look at changes in age patterns of mortality over time and thereby generate an independent estimate of deaths due to AIDS. This approach has been used in Zimbabwe[28] and South Africa,[29] where the results have helped to improve estimates of HIV/AIDS. Estimates of orphan numbers in countries with generalised epidemics derived from sentinel surveillance of HIV[30] can be compared with estimates from household questionnaires to further help validate UNAIDS/WHO estimates of HIV prevalence and AIDS mortality.[31] In countries with concentrated epidemics orphan data are less useful as checks on HIV/AIDS estimates because AIDS orphans represent only a small fraction of all orphans. However, in some of these countries case reports of HIV, AIDS, or AIDS deaths can help set bounds on the estimates.

In countries with generalised epidemics, national population based surveys are increasingly becoming available, providing an additional sources of data on HIV prevalence. Population prevalence surveys typically provide HIV estimates for both sexes and also cover more remote rural areas. Available population survey data have been used to help refine the estimates for several countries in the UNAIDS/WHO 2003 estimates. However, it must be noted that although the sampling frame in population prevalence surveys is designed to reach a nationally representative sample, absence from the household and unwillingness to be tested can bias the results.[32] Appropriate adjustments have therefore been made to take this bias into account, and additional work is ongoing to further inform these adjustments.

Final conclusions and future directions

In our approach to estimating plausible bounds around estimates of HIV/AIDS we have tried to include all major sources of uncertainty. However, some sources of uncertainty are difficult to estimate and have not been included. These include the availability of antiretroviral therapy (ART), survival of those on ART and effectiveness of prevention of MTCT programmes. Although information on coverage and effectiveness of these programmes is taken into account when making estimates of HIV, the uncertainty around the assumptions is not included in producing the plausibility bounds. For the UNAIDS/WHO 2003 estimates the current low coverage levels for most low and middle income countries makes this source of uncertainty quite small.[33] However, as coverage increases this issue will become more important, especially in estimates of new infections and AIDS mortality. Additionally we have not considered errors resulting from problems with laboratory testing. While there are numerous studies showing that both sensitivity and specificity of the commonly used HIV tests are very high,[34] these values do not capture the effects of poor sample handling and storage, or the use of out of date tests and poor reagents. Although quality assurance programmes reduce the likelihood of this source of error, this is not always the case. For example, the estimate of 33.7% adult HIV prevalence in Zimbabwe for the end of 2001 was based largely on surveillance data from the year 2000. An unrealistic drop in prevalence for subsequent years led to retesting of the samples for the year 2000 using a more specific test kit. This suggested that problems with laboratory error had resulted in an upward bias in prevalence of about 14% in the 2001 estimate.[35]

Global infectious disease statistics are rarely presented with bounds based on any formal or rigorous procedure. This

leads to unnecessary debate about different estimates for specific diseases when in fact the difference may not be significant. We have presented methods to estimate plausible bounds about the UNAIDS/WHO estimates for HIV/AIDS at the end of 2003. This work represents a first step towards providing global disease statistics with more rigorously derived plausibility bounds. Future work will be needed to refine these methods.

. . . .

Authors' affiliations

N C Grassly, G Garnett, Department of Infectious Disease Epidemiology, Imperial College London, Norfolk Place, London W2 1PG, UK
M Morgan, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA 30333, USA
N Walker, UNICEF, 3 United Nations Plaza, New York, USA
K A Stanecki, P D Ghys, UNAIDS, 20 Avenue Appia, CH-1211 Geneva 27, Switzerland
J Stover, The Futures Group, Glastonbury, Connecticut, USA
T Brown, The East-West Center, Honolulu, Hawaii, USA

REFERENCES

1 **The United Nations Joint Programme of HIV/AIDS (UNAIDS)**. *Report on the Global HIV/AIDS Epidemic, June 1998*. Geneva: UNAIDS, 1998.
2 **The United Nations Joint Programme of HIV/AIDS (UNAIDS)**. *Report on the Global HIV/AIDS Epidemic, June 2000*. Geneva: UNAIDS, 2000.
3 **The United Nations Joint Programme of HIV/AIDS (UNAIDS)**. *Report on the Global HIV/AIDS Epidemic, 2002*. Geneva: UNAIDS, 2002.
4 **The United Nations Joint Programme of HIV/AIDS (UNAIDS)**. *Report on the Global HIV/AIDS Epidemic, 2004*. Geneva: UNAIDS, 2004.
5 **Walker N**, Stover J, Stanecki K, et al. The workbook approach to making estimates and projecting future scenarios of HIV/AIDS in countries with low level and concentrated epidemics. *Sex Transm Infect* 2004;**80**(Suppl I):i10–3.
6 **Ghys PD**, Brown T, Grassly NC, et al. The UNAIDS Estimation and Projection Package: a software package to estimate and project national HIV epidemics. *Sex Transm Inf* 2004;**80**(Suppl I):i5–9.
7 **Stover J**. Projecting the demographic consequences of adult HIV prevalence trends: the Spectrum Projection Package. *Sex Transm Inf* 2004;**80**(Suppl I): i14–8.
8 **Kilian AHD**, Gregson S, Ndyanabangi B, et al. Reductions in risk behaviour provide the most consistent explanation for declining HIV-1 prevalence in Uganda. *AIDS* 1999;**13**:391–8.
9 **Fylkesnes K**, Ndhlovu Z, Kasumba K, et al. Studying dynamics of the HIV epidemic: population-based data compared with sentinel surveillance in Zambia. *AIDS* 1998;**12**:1227–34.
10 **Fylkesnes K**, Musonda RM, Sichone M, et al. Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population based surveys. *AIDS* 2001;**15**:907–16.
11 **Glynn JR**, Buve A, Carael M, et al. Factors influencing the difference in HIV prevalence between antenatal clinic and general population in sub-Saharan Africa. *AIDS* 2001;**15**:1717–25.
12 **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**:849–55.
13 **Gregson S**, Terceira N, Kakowa M, et al. Study of bias in antenatal clinic HIV-1 surveillance data in a high contraceptive prevalence population in sub-Saharan Africa. *AIDS* 2002;**16**:643–652.
14 **Kwesigabo G**, Killewo JZ, Urassa W, et al. Monitoring of HIV-1 infection prevalence and trends in the general population using pregnant women as a sentinel population: 9 years experience from the Kagera region of Tanzania. *J Acquir Immune Defic Syndr* 2000;**23**:410–417.
15 **Programme National de Lutte Contre Le SIDA**. *Population based serosurvey*. Rwanda: Ministry of Health, 1998.
16 **Crampin AC**, Glynn JR, Ngwira BM, et al. Trends and measurement of HIV prevalence in northern Malawi. *AIDS* 1999;**17**:1817–25.
17 **UNAIDS Reference Group**. Improved methods and assumptions for estimation of the HIV/AIDS epidemic and its impact: recommendations of the UNAIDS Reference Group and Estimates, Modelling and Projections. *AIDS* 2002;**16**:W1–16.
18 **Collaborative Group on AIDS Incubation and Survival including the CASCADE EU Concerted Action**. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. *Lancet* 2000;**355**:1131–1137.
19 **Wilks SS**. The large-sample distribution of the likelihood-ratio for testing composite hypotheses. *Ann Math Stat* 1938;**9**:60–62.
20 **Walker N**, Garcia-Calleja JM, Heaton L, et al. Epidemiological analysis of the quality of HIV sero-surveillance in the world: how well do we track the epidemic? *AIDS* 2001;**15**:1545–54.
21 **National AIDS and STD Control Programme**. *National HIV prevalence in Kenya*. Nairobi: Ministry of Health, 2003.
22 **National AIDS Commission**. *Estimating national HIV prevalence in Malawi from sentinel surveillance data: technical report*. Lilongwe: Ministry of Health and Population, 2003.
23 **Bickel PJ**, Doksum KA. *Mathematical statistics: basic ideas and selected topics, volume 1*. London: Prentice Hall, 2001.
24 **Klir GJ**. Generalized information theory. *Fuzzy Sets and Systems* 1991;**40**:127–42.
25 **Friedman N**, Halpern JY. Plausibility measures: a user's manual. In *proceedings of Uncertainty in Artificial Intelligence Conference*. UAI '95, 1995:175–84.
26 **Habicht JP**, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology* 1999;**28**:10–18.
27 **Victora CG**, Habicht JP, Bryce JW. Evidence-Based Public Health: Moving Beyond Randomized Trials. *American Journal of Public Health* 2004;**94**:400–5.
28 **Feeney G**. The impact of HIV/AIDS on adult mortality in Zimbabwe. *Population and Development Review* 2001;**27**:771–780.
29 **Timæus IM**, Dorrington R, Bradshaw D, et al. Adult Mortality in South Africa, 1980–2000: From Apartheid to AIDS, in Population Association of America 2001 Annual Meeting Final Program and Abstracts, Grand Hyatt Hotel, Washington, D.C., March 29–31. MD: Silver Spring, 2001.
30 **USAID, UNAIDS, UNICEF**. *Children on the Brink 2002*. Washington, DC: Synergy Project, 2002.
31 **Grassly NC**, Lewis JJC, Mahy M, et al. Comparison of survey estimates with UNAIDS/WHO projections of mortality and orphan numbers in sub-Saharan Africa. *Popul Stud* 2004;**58**:in press.
32 **Boerma JT**, Ghys PD, Walker N. Estimates of HIV-1 prevalence from national population-based surveys as a new gold standard. *Lancet* 2003;**362**:1929–31.
33 **WHO**. *Coverage of AVT and PMTCT*. Geneva: WHO, 2003.
34 **WHO/UNAIDS**. *Operational characteristics of commercially available assays to determine antibodies to HIV-1 and/or HIV-2 in human sera. Report 11*. Geneva: WHO/UNAIDS, 1999.
35 **National AIDS Programme of Zimbabwe**. *National survey of HIV and syphilis prevalence among women attending antenatal clinics in Zimbabwe 2001: A comparative analysis with 2000 survey results*. Harare: Ministry of Health and Child Welfare, Department of Disease Prevention and Control, 2002.