Estimates of HIV incidence from household-based prevalence surveys

Timothy B. Hallett[a], John Stover[b], Vinod Mishra[c], Peter D. Ghys[d], Simon Gregson[a,e] and Ties Boerma[f]

Objective: To estimate HIV incidence in the general population in countries where there have been two recent household-based HIV prevalence surveys (the Dominican Republic, Mali, Niger, Tanzania, and Zambia).

Methods: We applied a validated method to estimate HIV incidence using HIV prevalence measurement in two surveys.

Results: We estimate incidence among men and women aged 15–44 years to be: 0.5/1000 person-years at risk in the Dominican Republic 2002–2007, 1.1/1000 in Mali 2001–2006, 0.6/1000 in Niger 2002–2006, 3.4/1000 in Tanzania 2004–2008, and 11.2/1000 in Zambia 2002–2007. The groups most at risk in these epidemics are typically 15–24-year-old women and 25–39-year-old men. Incidence appears to have declined in recent years in all countries, but only significantly among men in the Dominican Republic and Tanzania and women in Zambia.

Conclusion: Using prevalence measurements to estimate incidence reveals the current level and age distribution of new infections and the trajectory of the HIV epidemic. This information is more useful than prevalence data alone and should be used to help determine priorities for interventions.

© 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins

Introduction

HIV surveillance primarily uses measurements of HIV prevalence, although this records the historic, rather than current, trajectory of the epidemic, and high prevalence among subpopulations only indicates periods of risk at prior ages. Direct measurements of HIV incidence would be better for monitoring the course of the epidemic and identifying those most at risk, but this requires follow-up of large cohorts. Cohort studies [1,2] are expensive and rare, and the cohorts may not be representative of the country as a whole. Indirect measurements of incidence, using assays that discriminate new infections in cross-sectional surveys, have so far proved unreliable [3,4].

In recent years, an increasing number of household-based surveys [in particular, the Demographic and Health Surveys (DHS) [5] have tested participants for HIV,

generating prevalence measurements that are broadly representative of the general population [6]. These cross-sectional surveys are typically repeated in countries at 5-year intervals, and some countries have already had two surveys with HIV testing (and results from others are due shortly). Previously, we developed a method that uses the HIV prevalence measurements in two surveys to derive estimates of HIV incidence in the intervening period [7]. The method was tested using data from community-based cohorts, and the estimates were found to be in good agreement with direct measurements of incidence [7]. In this concise communication, we apply our method for estimating incidence to countries where there have been two household-based surveys with HIV testing and assess whether there is an evidence of changes in HIV incidence during the past decade.

Methods

Our method for estimating incidence from two cross-sectional measurements of prevalence has been published elsewhere [7]. Briefly, it is based on the synthetic cohort principle, whereby we assume that individuals of age a years in the first survey will be represented by individuals aged $a + \tau$ years in the second survey, in which τ is the interval between surveys, even though the surveys do not include the same individuals. The change in HIV prevalence among individuals aged a years in the first survey and $a + \tau$ years in the second survey can be attributed to incident infections and AIDS deaths. By finding an approximate value for the rate of AIDS deaths, based on the observed distribution of survival after infection [8] ('Method 2' in our earlier publication [7]), HIV incidence for that age group can be estimated. The effects of the increasing availability of antiretroviral therapy are accounted for in the calculation. We use simple bootstrapping (10 000 iterations) to quantify the error in the estimates of HIV incidence due to sampling errors in the measurement of prevalence [9]. Further information and a spreadsheet implementation of the method are available in the technical appendix.

The countries where there have been two household-based surveys with HIV testing (and sample sizes for men and women for dataset used) are: the Dominican Republic [2002 ($n = 10\,708$ and $10\,732$) and 2007 ($n = 23\,513$ and $24\,564$)], Niger [2002 ($n = 2988$ and 3062) and 2006 ($n = 2822$ and 4374)], Zambia [2001/2002 ($n = 1734$ and 2072) and 2007 ($n = 4941$ and 5502)], Mali [2001 ($n = 2596$ and 3835) and 2006 ($n = 3613$ and 4527)], and Tanzania [2003/2004 ($n = 4995$ and 5752) and 2007/2008 ($n = 6760$ and 8077)]. All of these are standard DHS [5], with the exception of Niger 2002 [10] and Tanzania 2007/2008 [11], where slightly different survey methods were followed. The age ranges covered by the surveys varied, but all included men and women aged 15−49 years in both rounds, allowing a common measurement of incidence among 15−44-year-olds. In the calculation, the interval between the surveys was taken as the difference between the midpoints of the fieldwork periods, rounded to the nearest integer.

HIV prevalence data from two surveys provide an estimate of the incidence rate for the inter-survey period. To assess a temporal trend in incidence, a third survey would be required, but no country has yet completed three national surveys. An indication of the temporal trend in incidence is made by comparing the estimates of incidence in the inter-survey period with those relating to the period before the first survey, which is calculated by assuming that prevalence was constant in the 5 years preceding the first survey [12]. This assumption was checked by inspection of trends in national prevalence estimates [13] (Fig. S1).

Results

Figure 1 shows the estimates of HIV incidence by age and sex, and Table 1 provides summary estimates for 15−44-year-olds.

Incidence among 15−24-year-old women is greater than that among men of the same age in Niger, Tanzania, and Zambia; the difference is greatest in Tanzania, where young women face nine times the risk of infection of young men. In these countries, the incidence rate among women decreases sharply with age, and incidence among women aged 25−44 years is approximately half that among men of the same age. In Tanzania, Mali, the Dominican Republic, and Zambia, there is a nonsignificant trend towards higher incidence among women over 40 years compared with women at middle ages. In Mali, men are more likely to be infected than women at all ages, and this difference strengthens with increasing age.

In all settings except Niger, women face their greatest risk of infection before their 25th birthday, whereas men face the greatest risk of infection at middle ages (25−39 years). Broadly, the groups most at risk of infection in Tanzania and Zambia are 15−19-year-old women and 25−39-year-old men. The groups most at risk of infection in the smaller epidemics are: 20−29-year-old men (the Dominican Republic), 35−39-year-old men (Niger), and men of all ages and 15−19-year-old women (Mali).

Comparing the HIV incidence rates for the inter-survey period (roughly 2001/2002−2006/2007 for all countries, except Tanzania, where it refers to 2004−2008) with the 5-year period before the first survey suggests that HIV incidence has fallen in all five countries. Only in the Dominican Republic [incidence rate ratio (IRR) 0.44]
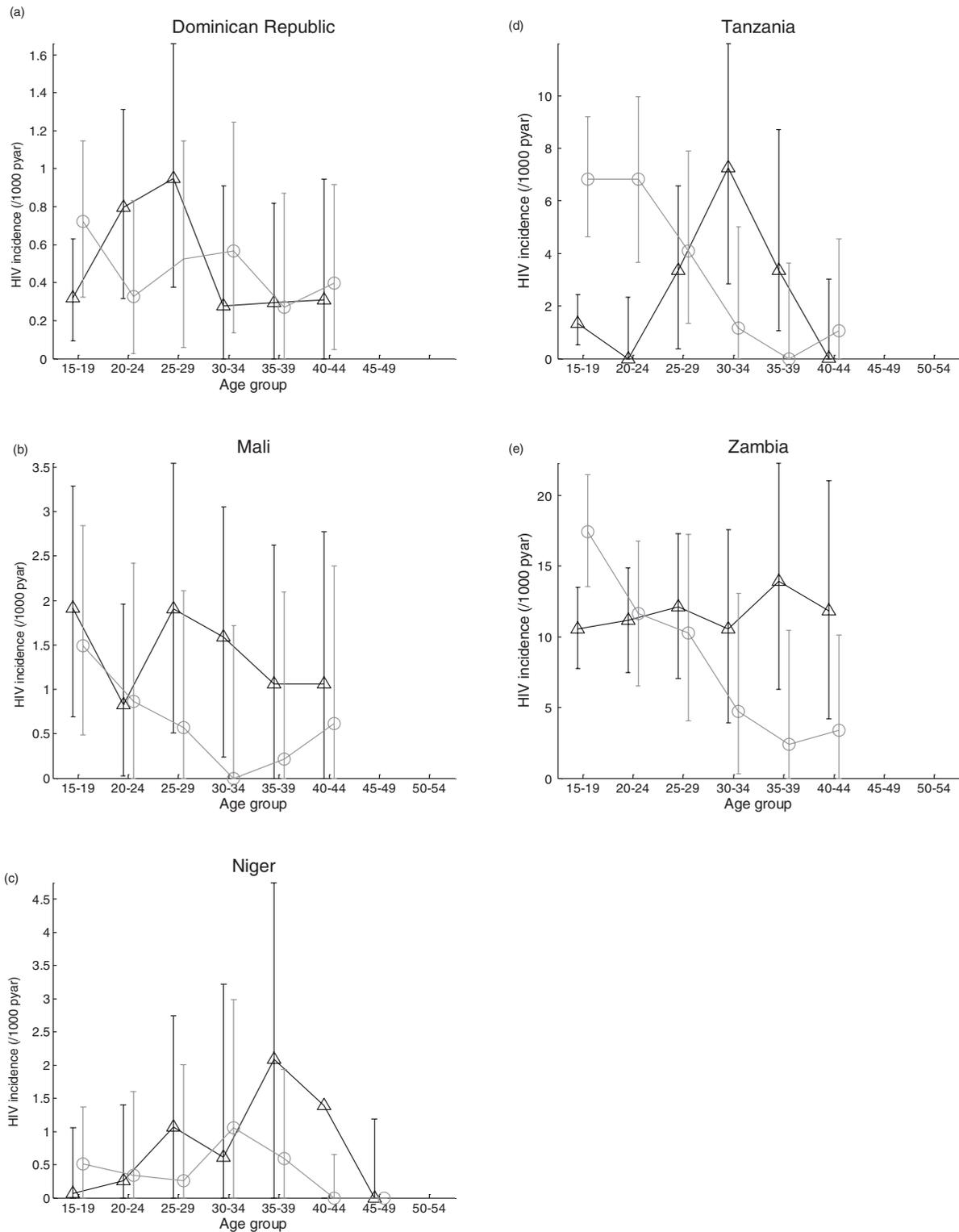
Fig. 1. Estimates of HIV incidence. Estimates are for men (triangles) and women (circles) per 1000 person-years at risk. Error bars shows 95% bootstrap intervals; bars approaching the zero line indicate that the observed change in prevalence can be completely ascribed due mortality. Note the differences in the scales of the vertical axes.

Table 1. Estimates of HIV incidence (15–44 years).

	Estimate (/1000 person-years)	95% Bootstrap interval		IRR (15–24 s women : men)[a]	IRR (25–44 s women : men)[b]	IRR (5 years)[c]
Low-prevalence countries						
Dominican Republic [0.8%] 2002–2007						
Overall	0.49	0.13	1.04	0.98	0.94	0.44
Men	0.50	0.15	1.05			0.38
Women	0.48	0.11	1.02			0.52
Mali [1.2%] 2001–2006						
Overall	1.14	0.23	2.65	0.85	0.23	0.59
Men	1.47	0.32	2.91			0.86
Women	0.78	0.13	2.37			0.36
Niger [0.7%] 2002–2006						
Overall	0.58	0.01	2.14	2.74	0.41	0.47
Men	0.69	0.02	2.49			0.62
Women	0.47	0.00	1.79			0.35
High-prevalence countries						
Tanzania [7.0%] 2004–2008						
Overall	3.37	1.52	6.42	9.33	0.50	0.44
Men	2.36	0.75	5.22			0.34
Women	4.42	2.32	7.65			0.53
Zambia [15.6%] 2002–2007						
Overall	11.22	6.60	16.52	1.38	0.52	0.64
Men	11.34	6.66	16.26			0.80
Women	11.07	6.54	16.81			0.53

Estimates are per 1000 person-years at risk and weighted by the number susceptible in the population. Values in square brackets show the HIV prevalence among adults of 15–49 years in the most recent survey. IRR, incidence rate ratio.
[a]IRR (15–24 s women : men) is the ratio of incidence among 15–24-year-old women in the intersurvey period, compared with men of the same age. [b]IRR (25–44 s women : men) is the ratio of incidence among 25–44-year-old women in the intersurvey period, compared with men of the same age. [c]IRR (5 years) is the ratio of incidence among 15–44-year-olds in the intersurvey period, compared with incidence in the 5 years before the first year (assuming constant prevalence). Values underlined indicate that the corresponding 95% bootstrap interval did not include 1.0.

and Tanzania (IRR 0.44) are the declines statistically significant overall (Fig. S4), and the reductions are greatest among men. In Zambia, there is evidence for incidence declines among women (IRR 0.53) (Table 1).

Discussion

Prevalence among adult women is typically higher than among adult men in large generalized epidemics [5], and our analysis reveals how this pattern is generated by very high rates of incidence among young women, despite much lower rates among older women. This pattern has also been observed in local community-based cohort studies [19] and is likely to be due to women being at most risk during early sexual experience, in particular through casual partnerships with older men [20,21]. The suggestion of an increase in incidence at older ages (>40 years) has also been identified in cohort studies [19], and it has been suggested that it is due to widowhood exposing women to the risk of infection as they form new partnerships [22]. Incidence among the youngest men in the high-prevalence settings is relatively low but then increases and remains at a higher level at older ages. The age/sex patterns of incidence in the lower prevalence

countries (Dominican Republic, Mali, and Niger) bear some similarities, but here, the most at-risk groups are generally middle-aged men not young women.

Our results also indicate substantial reductions in incidence during the current decade in all countries and significantly in the Dominican Republic and Tanzania, particularly among men, and among women in Zambia. However, it is not clear whether these changes are associated with changes in risk behavior or are only part of the natural evolution of the epidemic [23,24]. To make these calculations, we assumed that prevalence was constant (in each age group) in the 5 years before the first survey. This normally requires overall prevalence to have been stable for several years, which appears to be justified (Fig. S1), although there are some indications from antenatal clinic prevalence data that incidence gradually declined in the years before the first survey. If incidence did decline prior to the first survey, then our estimates of the change in incidence in the inter-survey period could be exaggerated. Trends in prevalence among 15–24-year-old women at antenatal clinics have been used as a proxy for trends in incidence [25,26]. In comparison, our approach for detecting changes in incidence can be related to all ages, and we have found previously that it is more likely to detect the full extent of changes in incidence [7].

Our estimates of incidence are associated with uncertainty arising from the sampling error in the prevalence estimates (range indicated by error bars in Fig. 1). Further, in the countries with smaller epidemics, household surveys could underestimate the true HIV prevalence, given the concentrations of infection in populations that are not fully captured (e.g. hostels, brothels, and military/police barracks) [27]. In all settings, nonresponse could also lead to underestimation of prevalence, although we expect this effect to be small and did not make a correction for it [27–29]. Nonetheless, if these biases are constant over time, trend information will be reliable, and despite errors in the absolute level of incidence, the estimated age distribution of incidence was similar across bootstrap iterations (Fig. S5).

As the incidence estimates at older ages are partly determined by the estimates for younger ages, the estimates of incidence for ages more than 40 years are the most uncertain (this factor is not reflected in the error bars). An additional source of uncertainty (not reflected in the error bars) is the set of assumptions made in the estimation method itself, including the distribution of survival with time since infection and the age distributions of those receiving treatment. As the age distribution of those on treatment is not recorded, we have assumed it follows a similar distribution to the expected AIDS deaths in the year of the survey in the absence of treatment (as projected in spectrum using time-series of the national estimates of prevalence [18]). However, this does not reflect how likely individuals are to access healthcare services. For instance, it may be that there are more young women on treatment than our assumption implies, as they are more likely to be have been tested for HIV (during pregnancy). To the extent that we have underestimated the number of individuals on treatment in a particular age group, the corresponding incidence estimate will be too high. However, as numbers on treatment are still small compared with the total number of people infected, and as treatment has not been widely available for long [13–16], resultant errors are likely to be modest. Our method also does not account for the possibility of different rates of international migration by HIV status and age, which could introduce bias. Such information on international migration is not available, but experimentation with the method suggests that the magnitude of any errors introduced by migration is likely to be small.

Although incidence measured in cohort studies remains the gold standard for accuracy, the communities studied are small in relation to the nation, and, after several years of intensive scientific study, may not be generally representative [1,2]. New assays, such as the BED test, which can detect recent infections have been used to estimate incidence from household-based sero-surveys [30,31], but due to the uncertain specificity of the test, results can be misleading [32,33]. The estimates of incidence presented here are based on a demographic method that makes a number of simplifying assumptions but which has been shown to reliably estimate incidence in other populations [7]. We, therefore, believe that these estimates will be of substantial value in monitoring the HIV epidemic in these and many other countries, as data from more surveys become available.
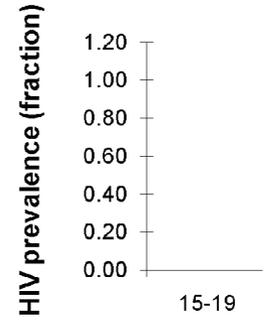
Acknowledgements

References

1. Zaba B, Whitworth J, Marston M, Nakiyingi J, Ruberantwari A, Urassa M, *et al*. **HIV and mortality of mothers and children: evidence from cohort studies in Uganda, Tanzania, and Malawi.** *Epidemiology* 2005; **16**:275–280.
2. Gregson S, Todd J, Zaba B. **Sexual behaviour change in countries with generalised HIV epidemics? Evidence from population-based cohort studies in sub-Saharan Africa.** *Sex Transm Infect* 2009; **85 (Suppl 1)**:i1–i2.
3. Parekh BS, Kennedy MS, Dobbs T, Pau CP, Byers R, Green T, *et al*. **Quantitative detection of increasing HIV type 1 antibodies after seroconversion: a simple assay for detecting recent HIV infection and estimating incidence.** *AIDS Res Hum Retroviruses* 2002; **18**:295–307.
4. UNAIDS Reference Group on Estimates Modelling and Projections. **Statement on the use of the BED-assay for the estimation of HIV-1 incidence for surveillance or epidemic monitoring.** *Wkly Epidemiol Rec* 2006; **81**:40–41.
5. Macro International Inc. **HIVPrevalence Estimates from the Demographic and Health Surveys.** http://www.measuredhs.com/topics/hiv/start.cfm. Calverton, Maryland, USA: Macro International, Inc.; 2008.
6. Boerma JT, Ghys PD, Walker N. **Estimates of HIV-1 prevalence from national population-based surveys as a new gold standard.** *Lancet* 2003; **362**:1929–1931.
7. Hallett TB, Zaba B, Todd J, Lopman B, Mwita W, Biraro S, *et al*. **Estimating incidence from prevalence in generalised HIV epidemics: methods and validation.** *PLoS Med* 2008; **5**:e80.
8. Todd J, Glynn JR, Marston M, Lutalo T, Biraro S, Mwita W, *et al*. **Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy.** *AIDS* 2007; **21**:S55–S63.
9. Kirkwood B, Sterne JAC. *Essential medical statistics*. 2nd ed. West Sussex, UK: Blackwell Science; 2003.
10. Louboutin-Croc J, Boisier P, Amadou Hamidou A, Oukem O. **Enquête Nationale de Séroprévalence de l'Infection par le VIH dans la population générale agée de 15 a 49 ans au Niger (2002).** Rapport final. In: *Centre de Recherche Médicale et Sanitaire*. www.cermes.ne; 2002.
11. National Bureau of Statistics Tanzania, MEASURE DHS. Tanzania HIV/AIDS and Malaria Indicator Survey 2007–08. Preliminary report; 2008.
12. Gregson S, Donnelly CA, Parker CG, Anderson RM. **Demographic approaches to the estimation of incidence of HIV-1 infection among adults from age-specific prevalence data in stable endemic conditions.** *AIDS* 1996; **10**:1689–1697.
13. UNAIDS. *Report on the global AIDS epidemic*. http://www.unaids.org/en/KnowledgeCentre/HIVData/GlobalReport/2008/2008_Global_report.asp. Geneva, Switzerland: UNAIDS; 2008.

14. Stover J. *Coverage of selected health interventions for HIV/AIDS prevention and care in less developed countries in 2001*. Geneva, Switzerland: World Health Organization; 2002.

15. Stover J. *Coverage of selected services for HIV/AIDS prevention, care, and support in low- and middle-income countries in 2003*. Washington, District of Columbia, USA: Constella Futures, POLICY Project; 2004.

16. Stover J, Fahnestock M. *Coverage of selected services for HIV/AIDS prevention, care, and treatment in low- and middle-income countries in 2005*. Washington, District of Columbia: Constella Futures, POLICY Project; 2006.

17. Egger M. **Outcomes of antiretroviral treatment in resource limited and industrialized countries.** *14th Conference on Retroviruses and Opportunistic Infections*. http://www.retroconference.org/2007/data/files/webpage_for_CROI.htm. Los Angeles, California, USA; 2007.

18. Stover J, Johnson P, Zaba B, Zwahlen M, Dabis F, Ekpini RE. **The Spectrum projection package: improvements in estimating mortality, ART needs, PMTCT impact and uncertainty bounds.** *Sex Transm Infect* 2008; **84 (Suppl 1)**:i24–i30.

19. Zaba B, Todd J, Biraro S, Shafer LA, Lutalo T, Ndyanabo A, *et al*. **Diverse age patterns of HIV incidence rates in Africa (TUAC0201).** *XVII International AIDS Conference*. Mexico City, Mexico; 3–8 August 2008.

20. Gregson S, Nyamukapa CA, Garnett GP, Mason PR, Zhuwau T, Carael M, *et al*. **Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe.** *Lancet* 2002; **359**:1896–1903.

21. Pettifor AE, Hudgens MG, Levandowski BA, Rees HV, Cohen MS. **Highly efficient HIV transmission to young women in South Africa.** *AIDS* 2007; **21**:861–865.

22. Lopman BA, Nyamukapa C, Hallett TB, Mushati P, Spark-du Preez N, Kurwa F, *et al*. **Role of widows in the heterosexual transmission of HIV in Manicaland, Zimbabwe, 1998–2003.** *Sex Transm Infect* 2009; **85 (Suppl 1)**:i41–i48.

23. Hallett T. **Monitoring HIV epidemics: declines in prevalence do not always mean good news.** *AIDS* 2009; **23**:131–132.

24. Hallett TB, Gregson S, Gonese E, Mugurungi O, Garnett GP. **Assessing evidence for behaviour change affecting the course of HIV epidemics: a new mathematical modelling approach and application to data from Zimbabwe.** *Epidemics* 2009; **1**:108–117.

25. World Health Organization/United Nations Joint Programme on AIDS. **Second generation surveillance for HIV: the next decade.** Geneva, Switzerland: WHO; 2000.

26. Zaba B, Boerma T, 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. Mishra V, Barrere B, Hong R, Khan S. **Evaluation of bias in HIV seroprevalence estimates from national household surveys.** *Sex Transm Infect* 2008; **84 (Suppl 1)**:i63–i70.

28. Marston M, Harriss K, Slaymaker E. **Nonresponse bias in estimates of HIV prevalence due to the mobility of absentees in national population-based surveys: a study of nine national surveys.** *Sex Transm Infect* 2008; **84 (Suppl 1)**:i71–i77.

29. Reniers G, Eaton J. **Refusal bias in HIV prevalence estimates from nationally representative seroprevalence surveys.** *AIDS* 2009; **23**:621–629.

30. Rehle T, Shisana O, Pillay V, Zuma K, Puren A, Parker W. **National HIV incidence measures: new insights into the South African epidemic.** *S Afr Med J* 2007; **97**:194–199.

31. Mermin J, Musinguzi J, Opio A, Kirungi W, Ekwaru JP, Hladik W, *et al*. **Risk factors for recent HIV infection in Uganda.** *JAMA* 2008; **300**:540–549.

32. Mermin J, Musinguzi J, Hladik W. **Estimating incidence of HIV infection in Uganda: reply.** *JAMA* 2009; **301**:160–161.

33. Barnighausen T, Wallrauch C, Welte A, McWalter TA, Mbizana N, Viljoen J, *et al*. **HIV incidence in rural South Africa: comparison of estimates from longitudinal surveillance and cross-sectional cBED assay testing.** *PLoS ONE* 2008; **3**:e3640.

Input: Time of Surveys (Midpoint)		
	Year	**Month Number for Estimate (1-12)**
First Survey		
Second Survey	*5.0*	*1*

Input: Prevalence Data	
	HIV Prevalence (Fraction)
Cohort's age group (*at first survey*)	**First Survey**
10-14	*0.00*
15-19	
20-24	
25-29	
30-34	
35-39	
40-44	
45-49	

Cohort's age group (*at second survey*)	**Second Survey**
15-19	
20-24	
25-29	
30-34	
35-39	
40-44	
45-49	
50-54	

In

HIV prevalence (fraction)

1.20
1.00
0.80
0.60
0.40
0.20
0.00

15-19

Input: Population Size	
Age-group	**Number**
15-19	
20-24	
25-29	
30-34	
35-39	
40-44	
45-49	
50-54	

Input: Numbers on Treatment		
Calendar Year	**Number**	**Month Number for Estimate (1-12)**
2001	*0*	
2002		
2003		
2004		
2005		
2006		
2007		
2008		
2009		
2010		
2011		
2012		