



National population-based HIV surveys – the method of choice for measuring the HIV epidemic

Thomas Rehle, Olive Shisana

Professor Thomas M Rehle is a Director and Senior Programme Advisor at the Human Sciences Research Council (HSRC). He also holds an appointment as Visiting Professor of International Health at the University of Cape Town, School of Public Health and Family Medicine. Professor Rehle is a member of the UNAIDS Reference Group on Estimates, Modelling and Projections and serves on the Steering Committee of the WHO Working Group on HIV Incidence Assays.

Dr Olive Shisana is Chief Executive Officer of the Human Sciences Research Council. She has served on many national and international scientific committees and advisory boards and on the US Institute of Medicine's Committee on Methodological Challenges in HIV Prevention Trials. She is currently serving on the Global HIV Prevention Working Group, the Emory University Global Health Institute Advisory Board and the South African National AIDS Council.

Corresponding author: O Shisana (oshisana@hsrc.ac.za)

National population-based surveys that include HIV testing are considered to be the 'gold standard' to measure HIV prevalence at a country level, since such surveys include men, non-pregnant women and children, and hence a much wider proportion of the population than do antenatal surveys.¹ Three national HIV household surveys have been conducted in South Africa, the first in 2002, then in 2005 and 2008. These collected data on HIV status, socio-demographic factors and behavioural determinants as well as exposure to prevention programmes which greatly enhanced the analysis and interpretation of the trends in HIV infection.²

In his appraisal of the 2008 survey report in this issue of SAMJ, Dorrington³ questions the reliability of the survey findings and expresses concerns about the validity of our conclusions. He compares the HIV prevalence data obtained in the surveys with the estimates produced by his model, the ASSA AIDS and Demographic Model.⁴ Findings that were not close enough to those projected by the ASSA model are declared 'not in line with expectation' and therefore implausible. In this debate one should not forget that it is



empirical data that greatly improved the outputs of projection models. A prime example of this process is the ASSA model: the ASSA estimates of HIV-infected people for the year 2005 dropped from 7.6 million to 5.2 million (!) after the creator of the model re-calibrated the projections with data available from the national HIV household surveys.^{5,6}

The 2008 survey report focused on the indicators selected for measuring progress in the implementation of the South African National Strategic Plan. In the plan for 2007 - 2011, the Human Sciences Research Council (HSRC) is requested to adjust the scope of analysis in order to accommodate this requirement. It is therefore surprising that Dorrington argues that it will be of little comfort to policy-makers.

The ASSA2003 model estimates of HIV prevalence by age correlate well with the results of the 2008 HSRC survey. However, Dorrington's interpretation that the increase in HIV prevalence from 2005 to 2008 in females 30 years and older is implausible needs to be addressed. Firstly, the prevalence among females 30 years and older was 13.6% in 2005 and 15.7% in 2008, an increase of 2.1%, which is more likely to be a result of increased access to antiretroviral treatment (ART) among women over 30 years than of an increase in HIV incidence. ART has increased the survival time of people living with HIV, and as a consequence HIV prevalence is likely to increase predominantly in the age groups more likely to need ART. This should be borne in mind when interpreting the 2008 findings on HIV prevalence.

Dorrington's assessment of the survey data suffers from serious inaccuracies. For example, in his comparison of HIV prevalence differences calculated from the 2002 and 2008 national household surveys and the 2002 and 2007 antenatal surveys respectively, he uses HSRC prevalence data from the age group 2 years and older instead of the appropriate 15 - 49 years age group. Table I shows the pattern of HIV prevalence trends in this age group by province for the corresponding survey periods (household surveys 2002 - 2008 v. antenatal surveys 2002 - 2007). The provincial prevalence trends observed in the national household surveys and the antenatal surveillance surveys were overall in good agreement, with 6 of the 9 provinces showing the same trends in both survey methodologies. On the national level, the difference in HIV prevalence between the 2002 and 2008 national household surveys was +1.3%, and the difference between the 2002 and 2007 antenatal surveys was +1.5%.^{7,8}

Trends over time are powerful tools to assess whether the observed changes are real, especially when the uncertainty around single survey estimates is brought into epidemiological context. Among children aged 2 - 14 years, a decline in HIV prevalence has been observed from 2002 to 2005 and from 2005 to 2008, and the difference between 2002 and 2008 reached statistical significance. But more important, the decline is also epidemiologically plausible. The increasing coverage of effective prevention of mother-to-child-transmission (PMTCT) programmes in the country has certainly shown its impact on the vertical transmission of HIV, with fewer HIV-infected infants moving each year into the 2 - 14-year age cohort. On the basis of this contextual evidence we argue that the observed decline in HIV prevalence was real. However, the increasing number of HIV-infected children receiving antiretroviral treatment will make interpretation of their HIV prevalence levels increasingly difficult in future.

HIV prevalence trends in teenagers aged 15 - 19 years and youth aged 20 - 24 years are shown in Table II. A decline in HIV prevalence was observed for both age groups from 2005 to 2008. In view of the reported substantive behavioural changes among youth between the 2005 and 2008 surveys we consider these declines in HIV prevalence plausible and more than just the result of statistical fluctuation. This is illustrated in Fig. 1, which shows reported condom use at last sex in the 15 - 24 age group in the three surveys. The increases in condom use at last sex were statistically significant in both male and female youth.

Dorrington dismissed the decline in HIV infection rates among teenagers aged 15 - 19 years who participated in the

Table I. HIV prevalence trends by province, 15 - 49-year age group: national household surveys v. antenatal surveys

Province	Household survey (2002 - 2008)	Antenatal survey (2002 - 2007)
KwaZulu-Natal	↑	↑
Mpumalanga	↑	↑
Free State	↓	↓
Gauteng	↓	↓
Northwest	↑	↑
Eastern Cape	↑	↑
Limpopo	↑	↑
Northern Cape	↑	↑
Western Cape	↓	↑
National	↑	↑

↑ represents an increase and ↓ a decrease in trend.

Table II. HIV prevalence among youth, South Africa 2002, 2005 and 2008

Age group (yrs)	2002			2005			2008		
	N	HIV %	95% CI	N	HIV %	95% CI	N	HIV %	95% CI
Youth (15 - 19)	1 143	5.9	4.0 - 8.8	2 154	5.9	4.3 - 8.0	1 928	4.4	3.0 - 6.5
Youth (20 - 24)	956	13.2	10.4 - 16.7	1 966	15.2	12.5 - 18.2	1 689	13.6	11.4 - 16.2

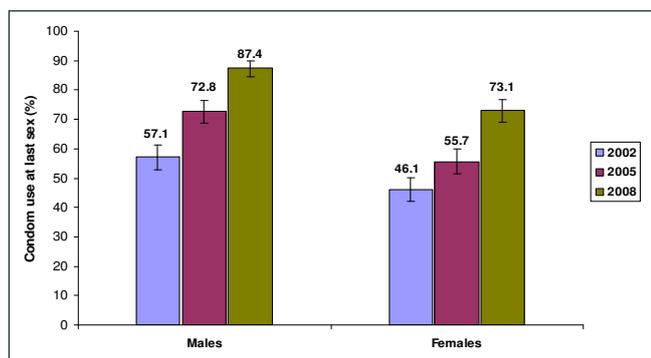


Fig. 1. Condom use at last sex, 15 - 24-year-olds.

2008 national HIV household survey as an unlikely result because similar declines were not observed in pregnant female teenagers attending public health clinics. This is like comparing apples with oranges. Clearly it stands to reason that pregnant teenagers are not representative of 15 - 19-year-old boys and girls in the general population. Teenage pregnancies are associated with a socio-demographic and behavioural risk profile that greatly increases the risk of HIV infection. This is shown in Table III, which compares HIV prevalence levels among the different teenage groups. The HIV infection rate in pregnant teenagers was 11.2%, over 2.5 times higher than the prevalence found in the whole group, 4.4% for males and females combined. However, it was close to the prevalence estimate for teenagers attending public antenatal clinics in 2007. Given the inherent risk profile associated with teenage pregnancy, we expect that the results will be similar in the 2008 antenatal survey.

Extrapolations from antenatal data to the general population should be made with caution. Antenatal sentinel data are subject to biases related to sampling, usage and coverage of selected antenatal services, differentials in risk behaviours and contraceptive use, and other socio-demographic factors (e.g. age distribution of pregnant women visiting antenatal clinics).¹ Taking into account the differential utilisation rate of public health services by race group in South Africa, Table IV compares HIV prevalence in all females as well as black females with pregnant females attending antenatal clinics in 2007. It is apparent that the HIV prevalence profile among black females participating in the 2008 national household survey is very similar to the 2007 antenatal survey, except among the young females.

HIV prevalence is the result of cumulative new infections (incidence) and cumulative deaths among HIV-infected persons

over time. This epidemiological paradigm provides the basis for methods that estimate HIV incidence from HIV prevalence data.^{9,10} Applying a mathematical approach, we have derived HIV incidence estimates from prevalence in young people aged 15 - 20 years (males and females combined) using prevalence data by single year of age and assuming that HIV prevalence differences between the age strata represent incident HIV infections. The change in the HIV incidence pattern was substantial for the 2008 survey year compared with the incidence figures calculated for the 2002 and 2005 survey years, especially for the single-year age groups 15, 16, 17, 18, and 19 years.² This straightforward method is best applicable in younger age groups when the effect of AIDS-related mortality on HIV prevalence levels is still minimal. We are currently extending the epidemiological HIV incidence estimation to the entire population 15 years and older, using a recently proposed method that infers population-level HIV incidence from prevalence obtained in two cross-sectional serosurveys.¹¹ The approach incorporates survival after infection and hence requires information of ART exposure in the study population. One of the novelties of the 2008 survey was the addition of ARV testing into the survey protocol, which will enable this type of analysis.

With all modelling approaches we should keep in mind the famous quote from George Edward Pelham Box, one of the greatest statisticians of the 20th century and a pioneer in the area of Bayesian inference: 'Essentially, all models are wrong, but some are useful.'¹² Substantial improvements have been made over the past years in modelling country-level HIV/AIDS epidemics. The UNAIDS Estimation and Projection Package (EPP), the tool of choice for preparing national estimates and projections in most of sub-Saharan Africa, has recently added uncertainty estimation in its projections for

Table IV. Comparison of HIV prevalence (%) among females by age group: national household survey 2008 v. antenatal survey 2007

Age group (yrs)	HSRC 2008 all females	HSRC 2008 black females	Antenatal survey 2007
15 - 19	6.7	7.8	12.9
20 - 24	21.1	24.3	28.1
25 - 29	32.7	38.9	37.9
30 - 34	29.1	36.2	40.2
35 - 39	24.8	33.0	33.2
40 - 49	15.3	22.2	21.5
Total	21.3	26.6	28.0

Table III. HIV prevalence (%) in 15 - 19-year age group: national household survey 2008 v. antenatal survey 2007

Age group (yrs)	Both sexes	Females	Pregnant females*	Antenatal survey 2007
15 - 19	4.4	6.7	11.2	12.9

*Currently pregnant or pregnant in the past 24 months as reported in the 2008 population-based survey.





generalised epidemics.¹³ We recommend that future versions of the ASSA model should also incorporate this capability of uncertainty analysis in order to present model estimates with 95% confidence bounds. This would greatly improve the comparison of ASSA outputs with empirical data from national population-based surveys.

1. Boerma TJ, Ghys PD, Walker N. Estimates of HIV-1 prevalence from national population-based surveys as a new gold standard. *Lancet* 2003; 362: 1929-1931.
2. Shisana O, Rehle T, Simbayi LC, et al. *South African National HIV Prevalence, HIV Incidence, Behavioural and Communications Survey*, 2008. Cape Town: Human Sciences Research Council Publishers, 2009.
3. Dorrington R. Does the 2008 HSRC survey indicate a turning tide of HIV prevalence in children, teenagers and the youth? (Debate). *S Afr Med J* 2009; 99: 631-633.
4. Actuarial Society of South Africa. ASSA2003 AIDS and Demographic Model. 2005. www.actuarialsociety.org.za/aids (accessed 30 November 2005).
5. Dorrington RE, Bradshaw D, Budlender D. *HIV/AIDS Profile of the Provinces of South Africa, Indicators for 2002*. Cape Town: Centre for Actuarial Research, Medical Research Council, Actuarial Society of South Africa, 2002.

Professor Dorrington responds:

Three major concerns about the analysis and reporting of the latest HSRC results were raised in my article, namely the potential for bias (given the low response rates), the lack of acknowledgement of uncertainty in the results, and the use of the results from the 2002 survey as the basis for implying trend. None of these issues has been dealt with in a satisfactory manner by Rehle and Shisana.

Their response does not address the question of bias. However, through providing survey results not published in the report, they inadvertently provide grounds for further concern on this issue with their comparisons in Tables III and IV, which purport to show the similarity between estimates from the survey and those from the 2007 antenatal survey (to which they assume the 2008 survey will be 'similar'). Bearing in mind the need for upward adjustment of the 2007 antenatal survey figures,^{1,2} both comparisons show the prevalence from the HSRC survey to be somewhat lower (2.7% lower than the correct figure for 2007^{1,2} in the case of the black women aged 15 - 49 years) than the figures with which the authors argue they should be comparable. Of course, given that probably around 90%³ of those tested in the national antenatal sample are black women, one must wonder why the authors chose not to use the prevalence among pregnant black women in their comparisons, as in past surveys (instead of pregnant women in Table III and all black women in Table IV).

The potential for bias is a crucial question deserving more debate. Arguments presented elsewhere by the authors (e.g. South African National AIDS Council and a UCT research seminar) that suggest that the survey is unbiased, either on the basis of research by Mishra and colleagues^{4,5} into this question with respect to household prevalence surveys carried out as part of the DHS surveys or on the basis of comparisons of the characteristics of the people who answered the questionnaire but did or did not agree to be tested, are problematic.

6. Dorrington RE, Johnson L, Bradshaw D, Daniel T. *The Demographic Impact of HIV/AIDS in South Africa. National and Provincial Indicators for 2006*. Cape Town: Centre for Actuarial Research, South African Medical Research Council and Actuarial Society of South Africa, 2006.
7. Department of Health. *Summary Report: National HIV and Syphilis Antenatal Sero-prevalence Survey in South Africa, 2002*. Pretoria: Department of Health, 2003.
8. Department of Health. *Summary Report: National HIV and Syphilis Antenatal Sero-prevalence Survey in South Africa, 2007*. Pretoria: Department of Health, 2008.
9. Saidel T, Sokal D, Rice J, Buzingo T, Hassig S. Validation of a method to estimate age-specific human immunodeficiency virus (HIV) incidence rates in developing countries using population-based seroprevalence data. *Am J Epidemiol* 1996; 144: 214-223.
10. Gregson S, Machekeano R, Donnelly CA, et al. Estimating HIV incidence from age-specific prevalence data: comparison with concurrent cohort estimates in a study of male factory workers, Harare, Zimbabwe. *AIDS* 1998; 12: 2049-2058.
11. Hallett TB, Zaba B, Todd J, et al. Estimating incidence from prevalence in generalised HIV epidemics: methods and validation. *PLoS Med* 2008; 5: e80.
12. Box GEP, Draper NR. *Empirical Model-Building and Response Surfaces*. New York: Wiley, 1987.
13. Brown T, Salomon JA, Alkema L, Raftery AE, Gouws E. Progress and challenges in modelling country-level HIV/AIDS epidemics: the UNAIDS Estimation and Projection Package 2007. *Sex Transm Infect* 2008; 84: Suppl 1, i5-i10.

To the criticism that it would be more useful and honest to acknowledge the uncertainty and publish the confidence intervals, the authors' response is that they regard 'epidemiological plausibility' as being 'more important' than statistical significance. They argue that the decrease in prevalence among children aged 2 - 14, from an implausible 5.6% in 2002 to a more sensible 2.5% in 2008, is 'real' based on the 'contextual evidence' that coverage of effective PMTCT programmes has increased. They present no quantitative evidence to explain how a programme preventing infection in infants and with low coverage between 2002 and 2005 might explain, in an 'epidemiologically plausible' way, that the bulk of the drop (5.6% to 3.3%) in prevalence among children aged 2 - 14 occurred between 2002 and 2005!

Similarly they argue that the drop in prevalence in the youth is plausible in the light of their 'reported substantive behavioural changes'. Ignoring the question about whether reported behaviour is actual behaviour, it is curious that the reported drop in knowledge (also statistically significant) in the youth, the increase in percentage of males with more than one partner in the past year, and the fall in the age of sexual debut of males are not mentioned as indicators of changes in sexual behaviour.

Furthermore, the argument that comparing the *trend* in prevalence in 15 - 19-year-olds as measured by the survey with the *trend* in the prevalence of 15 - 19-year-olds attending public antenatal clinics is 'like comparing apples and oranges' misses the point. If condoms are being used to prevent the spread of the disease and are the major source of contraception, then: (i) one would expect to see a change in the age distribution of women attending public antenatal clinics (which one doesn't see); and, more importantly, (ii) if the prevalence in young women is falling to the extent suggested by the report, then surely one would have expected to see the prevalence among pregnant women (as measured by the antenatal survey) falling



too (which it doesn't appear to be doing). Unless, of course, the suggestion is that prevalence is only falling in women who wouldn't have fallen pregnant had they had unprotected sex!

Finally, of the concern that 2002 is used as a basis for inferring trend the authors point out, quite correctly, that the change in prevalence from the HSRC survey shown in Table III of the article is that for the population aged 2 and older, whereas it would be more appropriate to consider the change in prevalence for the population aged 15 - 49. However, they fail to remedy this error, described by them as 'some serious inaccuracies', by providing the figures for women aged 15 - 49, preferring to argue that the trend implied by differencing the prevalence rates from the 2002 and 2008 surveys is 'in good agreement' with the trend from the antenatal surveys on the grounds that 6 of the 9 provinces showed change in the same direction. Aside from the fact that the chances of getting such a result or better are about 75% if one allocates the up and down arrows randomly, their comparison misses the point. It was the conclusion, based on the comparison from 2002 that prevalence had *dropped* in 4 provinces, which was at issue. The table with the correct figures is reproduced below (Table I). It is interesting

Table I. Difference in prevalence (%), 15 - 49 years, HSRC (2002 - 2008) v. antenatal surveys (2002 - 2007)

Province	HSRC survey (2002 - 2008)	Antenatal survey (2002 - 2007)*
Western Cape	-7.9	+2.9
Northern Cape	-0.6	+1.5
Free State	-0.9	+2.7
Gauteng	-5.1	-1.0

*These values ignore the impact of the expansion of the sample in 2006 which if allowed for would probably increase these differences by at least 1% and by as much as 3% for the Northern Cape in particular.

to note that not only do the corrected figures not change the argument, but in the case of 2 of the 4 provinces (Western Cape and Gauteng) the differences are even more marked.

Rehle and Shisana also argue that change in overall prevalence over the period of the two surveys is very similar. The problems with this argument are: (i) as mentioned in the footnote to Table I there was a significant change in the sample used by the antenatal survey in 2006 and this, if anything, probably leads to an underestimate of the trend between 2002 and 2007; and (ii) the prevalence in 2007 that is comparable to the 2002 figure is not 28.0% but 29.3%,^{1,2} and hence the implied increase in prevalence in women attending public antenatal clinics is at least 2.8% (which is a good deal higher than the 1.3% they report for the national household prevalence survey).

It should be noted that none of the points above have been argued on the basis of a model (ASSA's or otherwise). My purpose was not to argue that models are better than empirical data or that the HSRC survey is wrong (at least not in any way that suggested fault on the part of the investigators) or that PMTCT and ARV aren't having an effect or that behaviour is not changing towards the less risky, but to suggest that interpretation of the results should be more cautious and scientific and prepared to acknowledge the limitations of the survey.

1. Dorrington R, Bourne D. Has HIV prevalence peaked in South Africa? - Can the report on the latest antenatal survey be trusted to answer this question? *S Afr Med J* 2008; 98(10): 754-755.
2. Dorrington R, Bourne D. Re-estimated provincial HIV antenatal survey prevalence for 2007 and a reinterpretation of the national trend. *S Afr Med J* 2008; 98(12): 940-941.
3. Shisana O, Rehle T, Simbayi LC, et al. *South African National HIV Prevalence, HIV Incidence, Behavioural and Communications Survey, 2005*. Cape Town: HSRC, 2005.
4. Mishra V, Barrere B, Hong R, Khan S. Evaluation of bias in HIV seroprevalence estimates from national household surveys. *Sex Transm Inf* 2008; 84: i63-i70. http://sti.bmj.com/cgi/reprint/84/Suppl_1/i63.pdf (accessed 8 July 2009).
5. Mishra V, Hong R, Khan S, Gu Y, Liu L. *Evaluating HIV Estimates from National Population-Based Surveys for Bias Resulting from Non-Response*. DHS Analytical Studies No. 12. Calverton, Md: Macro International Inc., 2008.