Does the 2008 HSRC survey indicate a turning tide of HIV prevalence in children, teenagers and the youth?

Rob Dorrington

Recently the Human Sciences Research Council (HSRC) released its report on the results of its third and most recent household prevalence survey. The main conclusion of the analysis is that ‘some solid progress has been made in the fight against the disease in the past few years, especially among teenagers and children’. In particular the authors conclude:

- that HIV prevalence at national level has decreased among children aged 2 - 14 years, from 5.6% in 2002 to 2.5% in 2008
- that there was a substantial decrease in incidence in 2008 in comparison with 2002 and 2005, especially for the single age groups 15, 16, 17, 18 and 19 years
- that HIV prevalence has decreased among youth aged 15 - 24 years, from 10.3% in 2005 to 8.6% in 2008, and
- that HIV prevalence among adults aged 15 - 49 years has declined between 2002 and 2008 in the Western Cape, Gauteng, Northern Cape and the Free State, ‘with the largest decline of 7.9 [sic] percentage points in the Western Cape’?

How reliable are the results from the survey, and how reasonable are these conclusions?

Despite the fact that the overall response rate (not given in the report, but easily calculated as 53% for those aged 2+) is slightly lower than that reported for the previous survey (55%), the results of the survey appear to be as consistent with expectations as those of the previous survey. As shown in Fig. 1, overall, the prevalence (among those aged 2+) from the survey (10.9%) is 1% below that expected on the basis of projections of the ASSA2003 model (11.9%), although the prevalence for males (7.9%) is significantly lower than expected (10.8%).

There may be grounds for believing that the ASSA2003 model exaggerates the prevalence in men a little. However, as with the previous HSRC survey, if one calculates the prevalence directly using the age-specific prevalence rates given in the report (and estimates of the mid-year population) one gets estimates more in line with expectations – about 1% higher for males and 0.6% lower for females. This may suggest a problem with the weighting used in the survey to adjust for non-response. In the case of males in particular the lower prevalence could also be due in part to the fact that the survey may not be capturing some of the high-risk individuals. This could have happened either because they were expressly excluded (military, prisons, etc.) or because they are less likely than others to participate in such surveys (less likely to have a fixed home address or to be at home, etc.). In this regard it is a concern that the survey tested only about 46% of the number of men (aged 15+) expected given the number of valid visiting points.

Comparing the estimates of prevalence by age (Fig. 2) one can see that there is a fairly close correspondence between the results and what might be expected on the basis of the ASSA2003 model. This is especially true if one allows for lower participation of high-risk individuals. Nevertheless the model produces higher estimates of prevalence for young and old men and 20 - 24-year-old women, while the survey produces higher estimates than the model for older women. While it is quite possible that the model is underestimating the prevalence in older women, the prevalence rate for women aged 30+ from the survey has increased so significantly, from around 12.5% in 2005 to around 15.5% in 2008, that even although some increase is also apparent from the antenatal survey, the incidence necessary to support such a large increase in prevalence would be implausibly high.

Professor Dorrington is an actuary and demographer and Director of the Centre for Actuarial Research at the University of Cape Town. He has worked on estimating mortality in southern Africa for the past 20 years and extensively on modelling the demographic impact of HIV/AIDS for the past 15 years.

Corresponding author: R E Dorrington (rob.dorrington@uct.ac.za)
The correlation of provincial prevalence rates with those expected is less reliable than the subsequent surveys. For example, the 2002 survey is, for various reasons, considered to be eventually, in scientific journals. Unfortunately the report ignores (largely, and certainly in drawing its central conclusions) the uncertainty around the estimates, with the result that differences over time that are noted are often not statistically significant. A prime example of this is the conclusion that the incidence in teenagers and the youth is falling. Apart from several minor methodological concerns with the approach used to produce the estimates of incidence in teenagers, the uncertainty around these estimates, although not published, is likely to be as large as the estimates of prevalence themselves, and far exceed the differences between estimates from one age to the next! To see this we only have to consider the confidence intervals for teenagers (15 - 19 years) as a group. Although the report does not provide confidence intervals for males and females combined, it does so for each separately. For the girls the confidence intervals around the estimates were 7.1 - 12.4% for 2005 and 4.8 - 9.0% for 2008, and for boys 1.4 - 7.1% and 0.7 - 7.1%, respectively. As far as the youth as a whole (15 - 24 years) are concerned, the confidence intervals around the estimate of

As far as the analysis of the results goes the report suffers from two major problems. The first is that trends, in many cases, are asserted on the basis of a comparison of the results of the latest survey with those of the 2002 survey. Unfortunately the 2002 survey is, for various reasons, considered to be less reliable than the subsequent surveys. For example the correlation of provincial prevalence rates with those expected from the ASSA model was a very low 31% for 2002, while for 2005 and 2008 it is around 90%. Further, while the correlation of the estimates of provincial prevalence rates from the 2002 and 2005 HSRC surveys is barely 50%, that of the 2005 and 2008 surveys is 99%.

Using 2002 as the base year leads to some absurd conclusions. An example of this is the conclusion that a number of provinces have experienced substantial declines in prevalence over the recent past. The declines implied by the report are compared, in Table I, with the differences between the prevalence in 2007 and 2002 from the national antenatal survey. Clearly there is little support for the suggestion that these provinces actually experienced a substantial decline in prevalence.

A further example of the problems of using 2002 as the base year is the conclusion that prevalence in children aged 2 - 14 decreased from 5.6% in 2002 to 2.5% in 2008. While it is possible that prevalence has declined in children as a result of the roll-out of the prevention of mother-to-child-transmission (PMTCT) programme, the comparison with 2002 greatly exaggerates the extent of the decline, given the doubts about the accuracy of the improbably high estimate for 2002.

The second major problem is that the report ignores (largely, and certainly in drawing its central conclusions) the uncertainty around the estimates, with the result that differences over time that are noted are often not statistically significant. A prime example of this is the conclusion that the incidence in teenagers and the youth is falling. Apart from several minor methodological concerns with the approach used to produce the estimates of incidence in teenagers, the uncertainty around these estimates, although not published, is likely to be as large as the estimates of prevalence themselves, and far exceed the differences between estimates from one age to the next! To see this we only have to consider the confidence intervals for teenagers (15 - 19 years) as a group. Although the report does not provide confidence intervals for males and females combined, it does so for each separately. For the girls the confidence intervals around the estimates were 7.1 - 12.4% for 2005 and 4.8 - 9.0% for 2008, and for boys 1.4 - 7.1% and 0.7 - 7.1%, respectively. As far as the youth as a whole (15 - 24 years) are concerned, the confidence intervals around the estimate of

![Graph](image-url)

**Fig. 2. Age-specific prevalence rates from the 2008 HSRC survey compared with those expected from the ASSA2003 model.**

Unfortunately the report on the 2008 survey differs from previous reports in that it focuses on presenting results specific to the needs of monitoring the National Strategic Plan (NSP). As such it excludes results that are important for assessing the quality and consistency of the results of the survey and useful for other purposes. While there can be no disagreement that later peer-reviewed publication of further analysis will help ensure an acceptable standard of analysis, the undertaking to release further results in this manner is of little consolation to non-HSRC-consortium researchers and policymakers. It is not clear that the detail that is useful for researchers in general will always be publishable in scientific articles, and even if it is, non-consortium researchers and policymakers are forced to wait for key results (such as numbers tested by combinations of population group, province, sex and age group, and the prevalence of pregnant women attending public antenatal clinics by province and age group, etc.) to be published, eventually, in scientific journals.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Cape</td>
<td>-6.9</td>
<td>+2.9</td>
</tr>
<tr>
<td>Northern Cape</td>
<td>-2.5</td>
<td>+1.5</td>
</tr>
<tr>
<td>Free State</td>
<td>-2.3</td>
<td>+2.7</td>
</tr>
<tr>
<td>Gauteng</td>
<td>-4.4</td>
<td>-1.0</td>
</tr>
</tbody>
</table>

*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.

---


---

**SAMJ Forum**

September 2009, Vol. 99, No. 9 SAMJ
prevalence are 8.7 - 12.0% for the 2005 survey and 7.2 - 10.4% for the 2008 survey. In other words, the confidence intervals of estimates from the various surveys overlap to such an extent for it to be quite probable that the differences between these estimates are nothing more than random fluctuation.

Of course, none of this is to say that there has not been a fall in incidence at the younger ages. It says only that if there has been such a fall, the survey is much too small to detect it. And while the survey presents evidence of reported changes in behaviour that might support the possibility of falling incidence in youth, it is difficult to know how accurate responses to these questions are. Certainly, if there have been changes in behaviour they are not substantial enough to impact on the proportion of pregnancies to teenage mothers, as the proportion of teenagers attending public antenatal clinics has remained fairly constant at around 19 - 20% for a number of years and certainly from 2002 to 2007.10,11 The changes are therefore unlikely to have been substantial enough to impact significantly on the incidence in teenagers.

It will be interesting to see if the dramatically lower incidence among teenagers estimated in the report is supported in any way by the results of the 2008 national antenatal survey, hopefully to be released later this year. (Although it might appear from the antenatal survey results that there has been a decline in prevalence among teenagers since 2005, it is impossible to tell to what extent this is simply due to the significant change in sample from 2006. There is no evidence of a decline in prevalence in this age group from 2002 to 2005 or after 2006.)