

## Testing for Recent Infection to Estimate HIV Incidence from Single Cross-Sectional Surveys

Reshma Kassanjee - Statistical Researcher, SACEMA

### Promising Developments in Incidence Estimation

The first SACEMA Quarterly update of 2010 was devoted to the topic of HIV incidence (the rate of occurrence of new infections in a population). Incidence will always remain a crucial measure in epidemiology, providing a direct and current indication of the spread of disease. Prevalence (the fraction of the population with a condition at a point in time) is a metric more commonly measured, but is less informative as it emerges from historic incidence, demography and survival dynamics. South Africa has the largest HIV-positive population, exceeding 5 million individuals in 2011 (1). This year the government will launch its second 5-year National Strategic Plan (NSP) against HIV for 2012-2016. The headline goal of the first 5-year NSP (2007 – 2011) was to halve incidence, but it will be difficult to assess whether this goal has been achieved.

There are a number of approaches for estimating HIV incidence, with varying tractability, complexity and limitations. (i) In a prospective longitudinal study of a cohort of initially uninfected subjects, infection events are directly counted. However, such studies can be costly and prone to unrepresentative sampling. (ii) HIV prevalence, measured in sentinel or general populations, is often modelled to estimate incidence. For example, in age groups with low HIV mortality, differences in prevalence by age may be attributed to new infections. Alternatively, prevalence, measured at multiple time points in the past, could be used to estimate historic incidence. However, prevalence data has become increasingly complex to interpret as epidemics mature, and knowledge of post-infection mortality is often limited. (iii) Alternatively, back-calculation from AIDS cases involves estimating the historic HIV incidence that produced the observed AIDS incidence. However, this method provides little indication of recent incidence. An extension of this method utilises reported HIV diagnoses too. (iv) There are also a number of detailed 'microscopic' models, such as the UNAIDS Modes of Transmission model (see SACEMA Quarterly, March 2010), and dynamical models. These typically explicitly model the mechanisms of transmission of the virus through the population, requiring a number of quantitative assumptions.

Additionally, in recent years, there has been considerable interest in estimating HIV incidence from single cross-sectional surveys testing for 'recent infection' through laboratory-measured host or viral biomarkers (2). In a survey, the sizes of the

HIV-negative, 'recently infected' and 'non-recently infected' populations can be measured, and incidence estimated using knowledge of the dynamics of the 'recent infection' biomarker (3,4,5). Given the potential benefits arising from using single cross-sectional surveys for incidence estimation, this approach has been applied in numerous studies, and has caught the attention of prominent organisations worldwide. The World Health Organisation (WHO) Technical HIV Incidence Assay Working Group (HIVIWG) produced an extensive guide on the use of biomarkers for 'recent infection' for incidence estimation. The Centers for Disease Control and Prevention (CDC) continues to actively improve laboratory tests used to measure biomarkers that identify 'recent infection', and, earlier this year, supported the WHO's efforts by hosting the latest HIVIWG meeting (August 2011). Notably, the Bill and Melinda Gates Foundation (BMGF) awarded the Health Protection Agency (HPA) a grant to assess, compare and optimise recent infection tests. The group working on this three-year project (2011-2013) is called CEPHIA, the Consortium for the Evaluation and Performance of HIV Incidence Assays, and comprises HPA, Blood Systems Research Institute (BSRI); University of California, San Francisco (UCSF); and SACEMA.

Two key obstacles to cross-sectional biomarker-based incidence surveillance remain: the (i) lack of standardisation of terminology and methodology, and (ii) poor characteristics, and characterisation, of currently available tests.

### Characterising a Test for Recent Infection

Testing for recent infection for the purposes of incidence estimation differs, in key respects, from estimating how long each subject in a study has been infected. It is generally accepted that, in the context of incidence surveillance, there are two crucial characteristics of a test for recent infection (5,6).

Firstly, a mean duration of recent infection captures the average time spent 'recently infected'. To produce reliable incidence estimates, the times that seroconverters spend 'recently infected' should be sufficiently large so that an adequate number of subjects are observed in this state in a survey of feasible sample size. For example, testing for the absence of HIV-antibodies amongst those with detectable HIV viral loads in principle identifies recent infection. However, the transiency of this state implies that very few 'recently infected' subjects would be found in a cross-sectional survey.

Conversely, the 'recent infection' classification should not endure for extended periods of time post infection, as subjects who were infected long into the past will appear 'recently infected' in the survey, making incidence estimates less informative about current incidence.

Due to the substantial variability of the virus progression and immunological response, individual seroconverters can spend vastly different times classified as 'recently infected'. To capture that some seroconverters are classified as 'recently infected' long after infection, a second characteristic, termed the false-recent rate, has been introduced. This measures the proportion of individuals that appear 'recently infected' a long time post infection. This proportion needs to be small to limit uncertainty in incidence estimates.

For a test for recent infection to be potentially useful for incidence surveillance, a mean duration of recent infection of 4 to 12 months and a small false-recent rate, less than 2%, are considered acceptable (6). Crucially, to produce robust incidence estimates, the characteristics of the recent infection test must be well-known (7).

### **Seroconverting Blood Donors as a Resource for Characterising Recent Infection Tests**

In the past, estimation of recent infection test characteristics has relied on detailed longitudinal data. Specifically, data obtained from the regular follow-up of HIV-negative subjects, and then repeated testing for 'recent infection' amongst seroconverters, with small inter-test intervals, has been used. Such data is typically used to model the time each seroconverting subject spends 'recently infected', and this information is in turn used to estimate the average dynamic that is captured by the mean duration of recent infection. However, such data is expensive and logistically difficult to collect, and this has been an obstacle to the development of recent infection tests. Methods for using less well-characterised, but more easily captured, data could therefore greatly advance developments in this field. One such innovation is explored in the article 'Seroconverting Blood Donors as a Resource for Characterising and Optimising Recent Infection Testing Algorithms' (8), as briefly summarised below.

In the work, a readily-available source of specimens was identified, namely that of seroconverting blood donors. Utilising specimens from blood donors provides unique efficiencies as blood for transfusions is routinely collected and tested for HIV in most countries. In the study, repeat donors (in South Africa and the USA in the period 2001-2009) who were observed to seroconvert were tested for 'recent infection', using the specimens collected at the times of the first seropositive donations.

Data captured from such study designs has been overlooked in the past, because there is no follow-up of seroconverters and typically large intervals between HIV-tests (or donations). Such data

provides little detail at an individual subject level. A method of estimation that draws meaningful information from this data, about average biomarker dynamics, without focusing on individuals, was employed. The approach is based on maximising the likelihood of the overall set of observed 'recent' and 'non-recent' classifications. The technical details of the method are provided in the abovementioned article (8). For example, if all inter-donation intervals were fixed at one year, then (for each donor equally likely to have been infected at any time in the year between his/her HIV-negative and HIV-positive donation) the overall collection of classifications provides information about the average dynamics of the biomarker for one year post infection. In the work, this principle is extended to allow for varying inter-donation intervals.

While the estimates of the test characteristics obtained using this method are likely not sufficiently robust for the purposes of incidence estimation, the work demonstrates an approach to perform preliminary characterisations of tests for recent infection, using a readily available source of specimens. More precise data could subsequently be used to better characterise only the most promising tests.

In conclusion, the cost and difficulty in collecting very detailed data describing the dynamics of a biomarker testing for 'recent infection' has been an obstacle to the characterisation of these tests, and hence incidence estimation. The innovative preliminary characterisation of recent infection tests, utilising more easily-sourced data, is therefore a promising development in the field. Using tests for recent infection to estimate times since infection for individual subjects is potentially of great interest in public health – this is a fundamentally different approach and would require an appropriately modified way to characterise tests for recent infection. Of interest in this discussion is the application of biomarkers testing for 'recent infection' for incidence surveillance.

The ability to estimate incidence from a study performed at a single point in time offers great benefits, and therefore cross-sectional biomarker-based incidence estimation has drawn much interest in recent years. With the collective efforts of leading organisations and experts to address the remaining limitations in the field, there seems to lay a promise for many exciting developments in this area. Cross-sectional incidence estimation is certainly a topic that should be closely followed by epidemiologists, and could greatly advance population-level HIV incidence surveillance in years to come.

**Reshma Kassanje**, Statistical Researcher, SACEMA. Areas of interest: Statistical / mathematical applications in epidemiology, particularly relating to HIV incidence and tests for recent HIV infection. [rkassanje@sun.ac.za](mailto:rkassanje@sun.ac.za)

**References:**

1. Statistics South Africa. Mid-year population estimates 2009. Statistical release P0302. 2011. Pretoria: Statistics South Africa; 2009.
2. Busch MP, Pilcher CD, Mastro TD, et al. Beyond detuning: 10 years of progress and new challenges in the development and application of assays for HIV incidence estimation. *AIDS*. 2010;24(18):2763-2771.
3. human immunodeficiency virus incidence rates from a cross-sectional survey using early diagnostic tests. *Am J Epidemiol*. 1995;141(2):166–172.
4. McDougal JS, Parekh BS, Peterson ML, et al. Comparison of HIV type 1 incidence observed during longitudinal follow-up with incidence estimated by cross-sectional analysis using the BED capture enzyme immunoassay. *AIDS Res Hum Retroviruses*. 2006;22(10):945–952.
5. McWalter TA, Welte A. Relating recent infection prevalence to incidence with a sub-population of assay non-progressors. *J Math Biol*. 2010;60(5):687–710.
6. Incidence Assay Critical Path Working Group. More and better information to tackle HIV epidemics: Towards improved HIV incidence assays. *PLoS Med*. 2011;8(6): e1001045.
7. Welte A, McWalter TA, Laeyendecker O, Hallett TB. Using tests for recent infection to estimate incidence: Problems and prospects for HIV. *Euro Surveill*. 2010;15(24):pii:19589.
8. Kassanjee R, Welte A, McWalter TA, et al. Seroconverting blood donors as a resource for characterising and optimising recent infection testing algorithms for incidence estimation. *PLoS ONE*. 2011;6(6):e20027.