

# Testing bias in calculating HIV incidence from the Serologic Testing Algorithm for Recent HIV Seroconversion

Robert S. Remis and Robert W.H. Palmer

**Objective:** Incidence is critical in monitoring HIV infection in populations but often difficult to measure. The Serologic Testing Algorithm for Recent HIV Seroconversion (STARHS) can estimate HIV incidence from a single specimen at low cost. Nevertheless, HIV testing patterns may introduce bias, rendering interpretation of the STARHS result problematic. We found empirical evidence of such bias in Ontario using the STARHS formula with varied window periods

**Methods:** In a hypothetical population of homosexual men, we calculated HIV incidence from the STARHS assay on the basis of incidence density, study duration, STARHS window period and intertest interval. We also incorporated the increased likelihood of a newly infected person having an HIV test due to seroconversion illness or high-risk behaviours ('seroconversion effect' or SCE). We also varied the intertest interval inversely as a function of incidence density. To adjust incidence estimates for bias, we fit empirical STARHS data to an algebraic formula expressing measured HIV incidence as a function of SCE and incidence.

**Results:** Incidence density estimates were unbiased when SCE or incidence density–interval interactions were absent. However, estimated incidence density was higher than true incidence density in the presence of SCE, as much as seven-fold higher under certain conditions. The goodness-of-fit provided estimates with an excellent fit, yielding plausible results.

**Conclusion:** HIV incidence from STARHS may be strongly biased because of early testing in recently infected persons, resulting in substantial overestimation, at least amongst men who have sex with men. Thus, incidence estimates from STARHS must be interpreted with considerable caution. Nevertheless, incidence estimates may be adjusted to yield unbiased results.

© 2009 Wolters Kluwer Health | Lippincott Williams & Wilkins

*AIDS* 2009, **23**:493–503

**Keywords:** bias, HIV incidence, monitoring, Serologic Testing Algorithm for Recent HIV Seroconversion

## Introduction

Since the HIV pandemic began, monitoring trends in HIV incidence has been critical as this indicator reflects the rate of new infection and helps to evaluate prevention efforts. Sustained or increasing HIV incidence has been of serious concern in many populations and has driven new approaches to controlling the epidemic.

Despite its importance, HIV incidence is difficult to measure. Classically, one follows a cohort of uninfected persons to measure the rate of new infection. Cohort studies, however, are expensive, time-consuming and subject to bias introduced by selective recruitment and attrition, study effect and aging [1]. Other indirect approaches to determine incidence such as serial seroprevalence studies or seroprevalence amongst young

---

Ontario HIV Epidemiologic Monitoring Unit, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada. Correspondence to Robert S. Remis, MD, MPH, FRCPC, Dalla Lana School of Public Health, University of Toronto, Health Sciences Building, 155 College Street, Room 512, Toronto, Ontario M5T 3M7, Canada.

Tel: +1 416 946 3250; fax: +1 416 971 2704; e-mail: rs.remis@utoronto.ca

Received: 26 June 2008; revised: 24 October 2008; accepted: 14 November 2008.

DOI:10.1097/QAD.0b013e328323ad5f

persons may be useful but their interpretation can be problematic [2].

In July 1998, Janssen *et al.* [3] reported a remarkable new approach to measure HIV incidence. The technique, termed the Serological Testing Algorithm for Recent HIV Seroconversion (STARHS), uses a laboratory assay that differentiates recent from long-standing HIV infection. Janssen *et al.* described a 'detuned assay' based on the principle that HIV antibody titre is low early in infection and increases subsequently. If the time to develop the antibody response is characterized, one can calculate HIV incidence from a single specimen. This approach is efficient as only specimens confirmed positive by the diagnostic assay are tested; also, STARHS assays are generally inexpensive. This is in contrast to approaches in which specimens are tested for p24 antigen [4] or viral RNA [5] to identify recently infected persons in whom these markers are present and antibody is still undetectable. These approaches are expensive as both positive and negative specimens must be tested and imprecise when HIV incidence is low or the preseroconversion period is short.

Measuring HIV incidence from single specimens is extremely attractive and has been likened to alchemy, in which 'philosopher's stone' transforms lead into gold. Nevertheless, there are potentially serious problems in the application of STARHS, which have received limited attention to date. When incidence is determined from STARHS in specimens collected for diagnostic purposes, HIV test-seeking behaviour may bias incidence estimates. Persons may present for HIV testing due to factors associated with HIV infection such as symptoms of primary HIV infection. This has been termed 'acute retroviral syndrome' or 'seroconversion illness' and commonly presents as a flu-like illness 2–4 weeks following infection [6–8] and may be severe enough to require hospitalization [6,9]. Men who have sex with men (MSM), in particular, are aware of seroconversion illness and its significance and the occurrence of influenza-like illness may precipitate a decision to be tested for HIV earlier than may have otherwise [10,11]. Recently infected persons may also present soon after exposure because of anxiety related to the high-risk sexual behaviours or other exposures. Finally, patients recently infected by HIV may present because they acquire a symptomatic sexually transmitted infection (STI), leading them to seek treatment. We have named the phenomenon of earlier HIV testing 'seroconversion effect' (SCE).

In Ontario, we have been using STARHS since 1999 but were concerned that the resulting HIV incidence estimates may be difficult to interpret. We observed incidence densities substantially higher [12,13] than those from cohort studies in similar populations and higher than estimates derived from independent modelling

[14]. We therefore wished to better conceptualize and quantify the potential bias introduced by HIV testing behaviours.

In preparatory work, we empirically observed the impact of testing bias due to SCE. If HIV testing were independent of events related to infection, measured HIV incidence should be independent of the mean window period used in the calculation, that is, calculated HIV incidence would be the same at different mean window periods. Using data on testing from the Ontario HIV diagnostic service, we observed that HIV incidence was substantially lower using longer mean window periods. We also realized that it was theoretically possible to quantify the magnitude of this bias from the actual data. This is potentially important as it could provide the basis for adjusting for testing bias introduced by SCE.

## Methods

### Empirical evidence of seroconversion effect

To determine whether there is empirical evidence for SCE, we analysed data from the Ontario HIV Laboratory, which performs essentially all diagnostic HIV testing in Ontario [15]. We calculated HIV incidence varying the standardized optical density (SOD) cutoff values corresponding to a window period of 70–336 days using the basic formula described by Janssen *et al.* [3].

### Theoretical basis for the analytic approach

Though our study initially used a simulation approach [16], we later realized that an exact solution was possible and developed a formula to express the measured incidence as a function of several input parameters.

We began with the basic formula to calculate incidence [3]. We then modified this expression, not only incorporating terms that included the true incidence but also incorporating other parameters that may introduce bias. Our formula incorporated HIV incidence, study duration, window period of the detuned assay and the intertest interval. We were, in particular, interested in determining to what extent measured HIV incidence varied as a function of these parameters. The output was the bias factor, that is, the ratio of the measured incidence (i.e. by STARHS) to the true incidence input in the formula.

We incorporated two potentially important sources of bias related to HIV testing. The first is the SCE as described above. We also hypothesized that persons at higher risk of HIV infection may test for HIV more frequently. This could potentially increase HIV incidence determined through STARHS.

For the purpose of our analysis, we used the period of increased probability of an HIV test related to seroconversion illness or high-risk behaviour of 90 days, or about 3 months, following infection, as if such an event were to incite a person to be tested, it would likely be within this period.

### Derivation of the formula incorporating seroconversion effect

We developed an algebraic expression of the model parameters to calculate the estimated incidence in situations with nonuniform testing intervals and to evaluate the extent of bias in incidence estimates. We began with the basic formula used to calculate incidence from STARHS as described by Janssen *et al.*:

$$I_{\text{est}} = \frac{N_{\text{dis}}}{N_{\text{test}} T_{\text{win}}}$$

In which  $I_{\text{est}}$ , estimated (calculated) incidence;  $N_{\text{dis}}$ , number of discordant STARHS results (diagnostic assay + and STARHS assay -);  $N_{\text{test}}$ , total number of tests in the period of observation;  $T_{\text{win}}$ , mean window period of the detuned assay (time from diagnostic + to detuned +).

The derivation of the formula for the measured HIV incidence as a function of the true incidence and SCE is described in detail in the Appendix. It is expressed as follows:

For  $T_{\text{win}} \geq T_{\text{sce}}$

$$I'_{\text{est}} = \frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right) + \left(\left(N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})\right) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) P_{\text{sce}}}{\left(\left(\frac{N T_{\text{obs}}}{T_{\text{test}}}\right) + \left(N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})\right) P_{\text{sce}}\right) T_{\text{win}}}$$

and for  $T_{\text{win}} < T_{\text{sce}}$

$$I''_{\text{est}} = \frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right) + \left(\left(N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})\right) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) \left(\frac{P_{\text{sce}} T_{\text{win}}}{T_{\text{sce}}}\right)}{\left(\left(\frac{N T_{\text{obs}}}{T_{\text{test}}}\right) + \left(N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})\right) P_{\text{sce}}\right) T_{\text{win}}}$$

In which  $I_{\text{true}}$ , true incidence;  $I_{\text{est}}$ , calculated incidence;  $T_{\text{win}}$ , window period of the sensitive/less sensitive (S/LS) assay;  $N$ , number of persons tested;  $T_{\text{test}}$ , mean time between HIV tests (intertest interval);  $T_{\text{obs}}$ , study period (period under observation);  $T_{\text{sce}}$ , time period used to define  $P_{\text{sce}}$ ;  $P_{\text{sce}}$ , proportion of newly infected persons having 'additional' testing in  $T_{\text{sce}}$ .

All time period parameters ( $T$ ) are expressed in the same units (days or years).  $N$  represents the number of persons and  $I$  represents the incidence expressed as the number of new HIV infections per person per unit time defined as in  $T$ . It is assumed that  $I_{\text{true}}$  and  $T_{\text{test}}$  are constant (uniform) across  $T_{\text{obs}}$ ,  $T_{\text{obs}}$  is greater than  $T_{\text{win}}$  and  $P_{\text{sce}}$  is uniform across  $T_{\text{sce}}$ .

### Parameter values

The formula was evaluated for parameter values amongst MSM in Ontario, a province in Canada with intermediate levels of HIV incidence and prevalence. The initial calculations were carried out using base-case values for each of four parameters. We varied the values of these parameters over a range from minimum to maximum values for each combination of the four parameters. The parameter values (see Table 1 for base-case, minimum and maximum values of model parameters) and the methods used to derive their values were as follows.

#### Incidence

We carried out a literature search of cohort studies and modelled estimates of HIV incidence amongst MSM [14,17–20].

#### Study duration

We examined the range of likely scenarios on the basis of data availability and timeliness as well as reviewed a number of studies that have been carried out on diagnostic data.

#### Window period

We used the window period as determined by the US Centers for Disease Control using the Vironostika HIV-1 Microelisa System (Bio-Mérieux, Inc, Durham, North Carolina, USA) [21]. We used a base case of 133 days, which is the window period of the HIVAB HIV-1 EIA

(Abbott Laboratories, Abbott Park, Illinois, USA) used in the original detuned assay. The Vironostika assay allows the use of a variable window period from 70 to 336 days, each with a mean value for SOD established by the developers of the test.

#### Intertest interval

There is limited data on the mean intertest interval amongst groups at risk for HIV. We derived initial estimates by determining the estimated number of tests amongst MSM in Ontario [22] and the estimated number of MSM in Ontario [14]. We also carried out a literature search on testing patterns to determine whether our initial estimates were plausible.

**Table 1. Model parameters and values.**

	Units	Base case	Minimum	Maximum
Incidence ( $I_{\text{true}}$ )	/100 py	1.00	0.25	5.00
Study duration ( $T_{\text{obs}}$ )	Days	365	90	1095
Window period, detuned ( $T_{\text{win}}$ )	Days	133	70	336
Intertest interval, mean ( $T_{\text{test}}$ )	Days	1095	365	1825
			Medium	High
SCE (proportion $\leq 90$ days)	Proportion	0.00	0.20	0.40
$I_{\text{true}} - T_{\text{test}}$ association	Slope	None	Low	High

py, person-years; SCE, seroconversion effect.

We also wished to explore the potential impact of an association between HIV incidence and intertest interval. Persons at high risk for HIV infection may test more frequently, potentially introducing bias into the measurement of incidence. Although this may not always be the case, empirical data from the Argus study on MSM in Montreal revealed such an association (Lambert G., personal communication) as did a study from London, England [23]. We termed this ' $I_{\text{true}} - T_{\text{test}}$  interaction'. To determine the impact of such an association, we divided the population into five strata of incidence varying on a monotonic increasing scale and varied the intertest intervals inversely, that is, the strata with the highest incidence had the shortest intertest interval.

### Adjusting HIV incidence for seroconversion effect

The algebraic formula given above expresses the calculated HIV incidence as a function of true incidence and the SCE. Though we had no assurance that this formula would incorporate all potential sources of uncertainty, we observed from empirical data that the variation in HIV incidence observed as a function of mean window was expressed closely by our algebraic formula with fitted values of true incidence and  $P_{\text{SCE}}$ . Therefore, we developed a procedure using a goodness-of-fit approach to determine the values of true incidence and  $P_{\text{SCE}}$ , which generated values of incidence most precisely fitting the observed data. Using custom software written in APL+Win Version 4.0 (APL2000, Rockville, Maryland, USA), we varied the values of incidence over the range from zero to 20 per 100 person-years in increments of 0.01 and the value of the  $P_{\text{SCE}}$  from zero to 40% in increments of 1%. The values of incidence density and  $P_{\text{SCE}}$ , which generated a modelled curve of incidence as a function of window period, were varied to minimize the sum of the squares of the residuals between the modelled and observed incidence curve. The software to carry out the bias adjustment is available from the authors at [www.phs.utoronto.ca/ohemu](http://www.phs.utoronto.ca/ohemu). To test this application and adjust actual data, we used data from MSM and injection drug users (IDUs) tested in the diagnostic service of the Ontario HIV Laboratory to calculate HIV incidence at five different values of the window period from 97 to 244 days.

## Results

### Empirical evidence of seroconversion effect

Varying the SOD cutoff values equivalent to different window periods, we found that for MSM, and to a lesser extent for IDUs, the calculated HIV incidence was substantially lower when a longer mean window period was used. For MSM, calculated incidence was 2.28 per 100 person-years for a mean window period of 133 days; 1.69 per 100 person-years for 170 days and 1.16 per 100 person-years for 336 days. For IDUs, the calculated incidence for the same three window periods was 0.29, 0.26 and 0.21 per 100 person-years. These results provide empirical evidence for the occurrence of SCE.

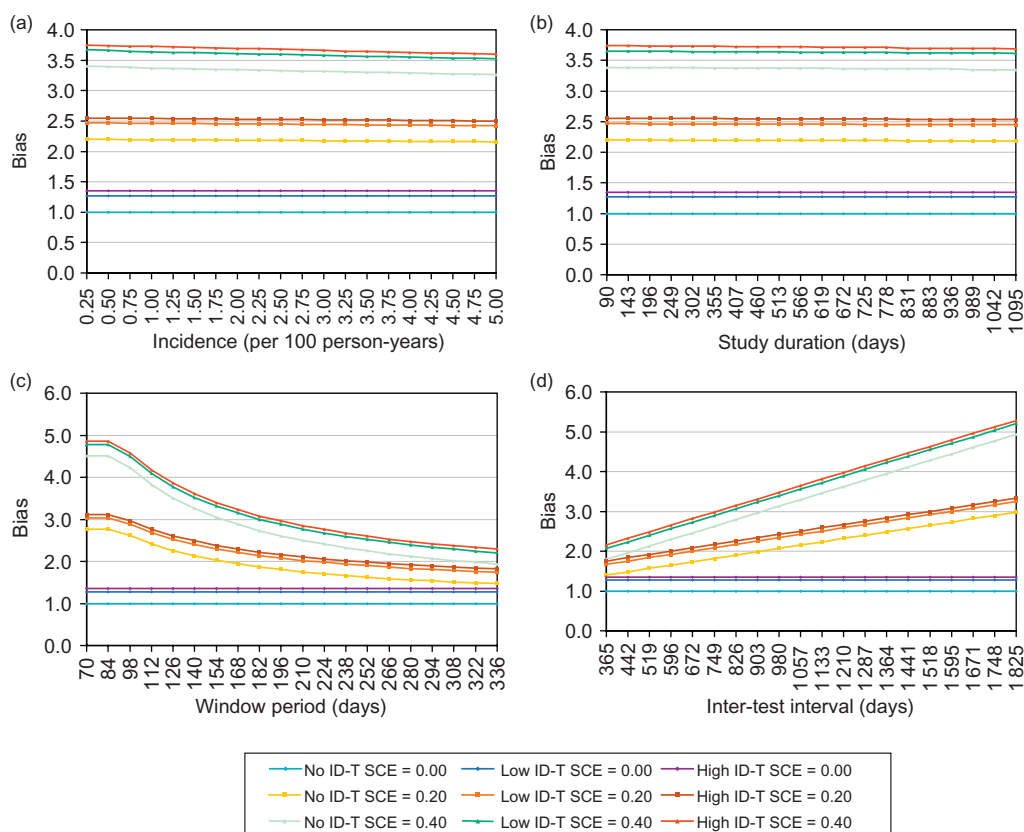
### Modelling the impact of testing bias

As indicated above, we quantified the bias as a ratio of calculated incidence to the true incidence over a wide range of combinations of the four parameters indicated above. For  $P_{\text{SCE}}$  equal to zero, and no interaction between incidence and intertest interval, we observed no bias. However, bias was present and, in some cases, to an important extent when  $P_{\text{SCE}}$  was not zero and  $I_{\text{true}} - T_{\text{test}}$  interaction was present. Table 2 shows a summary of the results for no and two levels of  $P_{\text{SCE}}$  and  $I_{\text{true}} - T_{\text{test}}$  interaction. Clearly, SCE was a greater source of bias than

**Table 2. Bias (calculated incidence/true incidence) as a function of incidence–intertest interaction and seroconversion effect.**

ID–T interaction	$P_{\text{SCE}}$		
	0.00	0.20	0.40
None			
Minimum	1.00	1.14	1.28
Mean	1.00	2.12	3.22
Maximum	1.00	3.96	6.90
Low			
Minimum	1.27	1.41	1.55
Mean	1.27	2.39	3.48
Maximum	1.27	4.23	7.15
High			
Minimum	1.35	1.49	1.63
Mean	1.35	2.47	3.57
Maximum	1.35	4.31	7.26

ID–T, incidence–intertest interaction;  $P_{\text{SCE}}$ , proportion with seroconversion effect.



**Fig. 1. Bias as a function of incidence–intertest interaction, seroconversion effect and (a) incidence, (b) study duration, (c) window period and (d) intertest interval. ID–T, incidence–intertest interaction; SCE, seroconversion effect.**

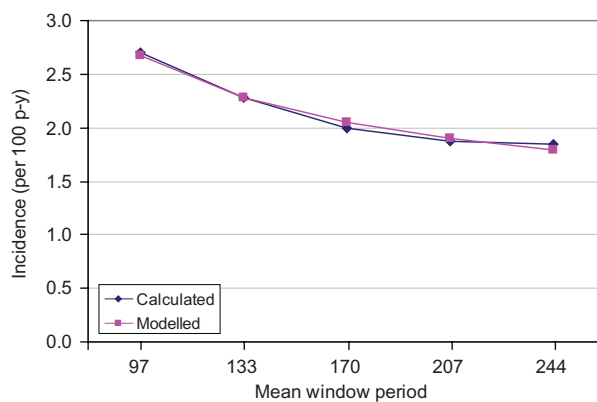
$I_{true} - T_{test}$  interaction. In the absence of SCE, the maximum bias was 1.35. However, in the presence of a high level of  $P_{SCE}$ , the mean and maximum biases were considerable, with bias as high as 7.26 at the highest tested level of  $P_{SCE}$  and  $I_{true} - T_{test}$  interaction. In general, in the presence of SCE, mean biases varied in the range of 2.12–3.57.

The bias observed over the range of values for each of the four parameters examined individually is summarized in Fig. 1(a–d). Figure 1a shows the effect of incidence and there appears to be a slightly decreased bias at increasing levels of incidence. There was no bias associated with study duration (see Fig. 1b). However, for the window period (see Fig. 1c), there was substantial bias, which was higher at shorter window periods. Finally, Fig. 1d displays the effect of varying intertest interval and its impact on bias; bias increased linearly and steeply with increasing intertest interval and the slope was dependent on the  $P_{SCE}$ .

### Adjusting HIV incidence for seroconversion effect

Figure 2 displays observed and modelled data using the adjustment procedure as described above. The observed curve represents incidence as a function of mean window period amongst MSM in Toronto using the original

formula provided by Janssen *et al.* [3] and the modelled curve represents incidence calculated using the more complex formula incorporating true incidence and  $P_{SCE}$  with the values of these parameters generated by our custom application. Table 3 presents some typical results of modelled and empirical values of incidence of density at different window periods for several exposure categories over the period 2001–2006. Thus, the iterative goodness-of-fit procedure achieved a relatively close fit to the empirical data for the exposure categories examined.



**Fig. 2. Calculated and modelled HIV incidence amongst men who have sex with men in Toronto, 2006.**

**Table 3. HIV incidence, unadjusted and adjusted by year for selected exposure categories, Toronto, 2001–2006.**

	2001	2002	2003	2004	2005	2006
<b>MSM</b>						
Unadjusted <sup>a</sup>	2.32	2.37	2.19	2.29	2.40	2.28
Adjusted	1.50	1.91	1.25	1.40	1.56	1.34
$P_{SCE}$ (%) <sup>b</sup>	8.2	3.9	10.9	9.5	6.9	9.9
$SS_{diff}$ <sup>c</sup>	0.0054	0.0233	0.0016	0.0072	0.0183	0.0063
<b>MSM-IDU</b>						
Unadjusted <sup>a</sup>	6.30	4.40	4.09	5.26	5.11	9.16
Adjusted	4.36	2.36	1.54	2.86	4.51	5.18
$P_{SCE}$ (%) <sup>b</sup>	8.4	13.4	23.6	12.9	3.4	13.9
$SS_{diff}$ <sup>c</sup>	0.6159	0.0488	0.0028	0.1986	0.4379	3.0275
<b>IDU</b>						
Unadjusted <sup>a</sup>	0.132	0.052	0.139	0.373	0.313	0.099
Adjusted	0.163	0.089	0.078	0.102	0.152	0.064
$P_{SCE}$ (%) <sup>b</sup>	0.0	0.0	10.8	36.1	16.4	7.9
$SS_{diff}$ <sup>c</sup>	0.0040	0.0038	0.0001	0.0000	0.0011	0.0001
<b>Heterosexual</b>						
Unadjusted <sup>a</sup>	0.029	0.025	0.031	0.033	0.023	0.022
Adjusted	0.021	0.019	0.025	0.029	0.015	0.026
$P_{SCE}$ (%) <sup>b</sup>	4.8	4.2	3.2	0.0	6.9	0.0
$SS_{diff}$ <sup>c</sup>	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

IDU, injection drug users; MSM, men who have sex with men; SCE, seroconversion effect.

<sup>a</sup>At SOD = 0.75 (mean window period 133 days).

<sup>b</sup> $P_{SCE}$  = proportion with SCE.

<sup>c</sup> $SS_{diff}$  = Difference in sum of squares between observed and fitted line.

## Discussion

In an analysis of potential biases in calculating HIV incidence from STARHS, we found that bias due to test-seeking behaviours was considerable. The values we used were plausible and based mostly on empirical data from MSM in Ontario. It is possible that biases may be different in other study populations though the range of values we used probably incorporated those from most other study populations. One limitation of our study is that we did not investigate the bias for different values of the duration of SCE, which, in our analysis, was taken as 90 days. For some parameters, good data are not readily available, especially for the interaction of incidence and testing frequency. Available data from both diagnostic services and special studies should be examined to further characterize these parameters. Clearly, however, anecdotal evidence and our results suggest that SCE is an important source of bias. A recent study [6] in Ontario found that a substantial majority of MSM who undergo HIV testing had symptoms suggestive of a primary HIV infection. It is unclear to what extent other populations at risk are aware of the occurrence of flu-like illness accompanying primary HIV infection. In some populations at risk or in special situations such as the blood donor setting in which symptoms and temperature are the basis for deferral, such behaviour could result in a bias in the opposite direction.

We incorporated two possible sources of bias into our model but there is a third possible source of bias, which might also be important in interpreting incidence estimates derived from STARHS. Participants included

in analyses to determine incidence may not be representative of the population for which the HIV incidence estimate is desired. In this regard, data collected from diagnostic testing databases may be particularly problematic. A good example is incidence based on STARHS using specimens collected amongst STI clinic patients. Thus, nonrepresentativeness introduces a further, potentially important, source of bias and must be taken into account in interpreting the results from STARHS.

To determine whether there was empirical evidence of bias, we reviewed three types of studies, which used STARHS to calculate HIV incidence, namely those using specimens collected for diagnostic purposes, epidemiologic studies and those examining blood donors. We hypothesized that bias would be present in the first group but to a lesser extent or not at all in the latter two groups.

Several studies [24–28] used diagnostic samples collected at STI clinics. In these studies, investigators observed very high incidence rates. A study by Weinstock *et al.* [24] in nine US STI clinics yielded an overall HIV incidence amongst MSM of 7.1 per 100 person-years. In Los Angeles, for example, the rate was 9.6 per 100 person-years as compared with a population-based estimate of 1 per 100 person-years at about the same time [20]. Such results are especially difficult to interpret because of the bias introduced by examining unrepresentative, high-risk persons.

Studies [29–31] from other diagnostic sites are also biased but probably to a lesser extent than those from STI clinics. Studies carried out at anonymous test centres from 2000 to 2001 in San Francisco yielded estimates of 2.5–3.3 per 100 person-years, about double the 1.4 per 100 person-years population-based estimate [20]. In Ontario, HIV incidence amongst MSM based on STARHS was 2.1 per 100 person-years [13]; this was considerably higher than our population-based estimate of one per 100 person-years [14].

We reviewed two epidemiologic studies; these would be less likely to show bias as we describe above since the HIV test is proposed by the investigators rather than sought by the participant. A study of IDUs by Kral *et al.* [32] observed an HIV incidence of 1.2 per 100 person-years in 1987–1998, actually slightly lower than the population-based estimate of 1.9 per 100 person-years [20]. Nevertheless, the 95% confidence limits were 0.7–2.0 per 100 person-years and thus included the latter estimate. Martindale *et al.* [33] used STARHS to calculate HIV incidence from baseline positive sera collected in a cohort study of MSM in Vancouver, Canada. Incorporating the STARHS data had no significant impact on overall HIV incidence in their cohort.

One might expect that incidence calculated from STARHS amongst blood donors would be relatively unbiased as the decision to donate blood is less likely associated with the risk of HIV infection (some persons

may donate blood to obtain an HIV test following a high-risk exposure but this is probably exceptional). In fact, studies [34,35] estimating HIV incidence amongst blood donors, both repeat and first time, yield very similar results to those from other methods [36]. Thus, as expected, there was little evidence of bias in this group. The results from STARHS are comparable to those from other methods based on seroconversion rates from repeat donors or RNA-yield rates from nucleic acid screening [37].

The US Centers for Disease Control recently published a study using back calculation and a novel approach using the STARHS result to estimate HIV incidence in the United States [38,39]. They concluded that HIV incidence was substantially higher than had been previously estimated, especially in men having sex with men and African-Americans. They examined the possibility that 'motivated testing' (akin to our 'SCE') might introduce bias into their results and lead to an overestimate of HIV incidence and concluded that the impact of such bias was about 7%. This is considerably lower than our empirical findings amongst MSM in Ontario. The difference may be due to the fact that their results were based on populations with different HIV testing patterns. Also, they based their estimate on self-reported reasons for HIV testing, which may not reliably capture the full impact of testing bias. A study [40] of MSM undergoing HIV testing in California revealed that men significantly downplayed risk behaviours that led them to be tested.

As incidence derived from STARHS is subject to important bias, it is useful to explore approaches through which this bias could be estimated and removed. Specific population-based studies could help us to better understand the patterns of HIV test-seeking behaviour and potentially estimate the values of  $P_{SCE}$  and  $I_{true} - T_{test}$  interaction. In particular, it would be useful to know the level of knowledge about seroconversion illness and the likelihood that a person would seek an HIV test because of symptoms compatible with primary HIV infection. Their effects could then be removed in the analysis.

In summary, studies reporting calculated incidence using STARHS, especially if based on specimens collected for diagnostic purposes, must be interpreted with caution because of the considerable bias introduced by HIV test-seeking behaviour. Our study revealed that bias due to early testing related to the SCE was important and could lead to marked overestimation of HIV incidence.

We also developed a procedure to remove the SCE. We believe that our approach provides a valid solution to control for this bias. The close fit of our algebraic formula expressing measured incidence as a function of true incidence and testing bias compared with empiric data over a range of mean window periods provides reassurance that our approach appropriately captures the functional relationship.

Our adjustment does not control for the association between testing frequency and HIV incidence. Nevertheless, in our analyses, we found this to be relatively unimportant. Also, the results we obtained can be generalized only to persons undergoing HIV testing. Most persons in highly affected populations undergo HIV testing though there may be some persons who do not undergo testing, either because they perceive themselves at low risk or are at high risk but fearful of being diagnosed HIV positive. It would be useful to explore the potential impact of nonrepresentativeness of the testing populations on HIV incidence calculated from STARHS.

## Acknowledgements

We wish to thank the following persons at the Ontario Ministry of Health and Long-Term Care as follows: Frank McGee, Coordinator, AIDS Bureau for core funding; Carol Swantee, Carol Major, Lynda Healey and Denise Soliven, HIV Laboratory for carrying out the detuned assay testing and Mark Vandenoort, Instructional Media Centre for graphic support. Janet Raboud provided helpful advice during the development of the methods. We would also like to thank Neil Hershfield, Vancouver for developing the software for adjusting incidence for bias and Juan Liu, University of Toronto for carrying out the bias adjustment calculations. The Laboratory Enhancement Study was funded by the Ontario HIV Treatment Network from 1999 to 2001 and the Public Health Agency of Canada from 2001 to 2008.

R.S.R. developed the theoretical conception of the biases inherent in the application of the STARHS assay in estimating HIV incidence and the general approach to modelling it. He also identified the empirical evidence for the presence of the SCE. R.W.H.P. developed the statistical model to assess bias under different assumptions of parameter values and derived the algebraic formula to express measured HIV incidence as a function of true incidence and SCE.

## Derivation of the formula

$I_{est}$	Estimated (calculated) incidence density
$N_{test}$	Mean number of tests in the period of observation
$N_{dis}$	Mean number of discordant tests [S(+) LS(-)]
$T_{win}$	Window period of the sensitive/less sensitive assay
$N$	Number of persons testing
$T_{test}$	Mean intertest interval
$T_{obs}$	Period of observation (study period)
$T_{win}$	Mean window period of the sensitive/less sensitive assay
$I_{true}$	True incidence density

Equation (1), below, has been used to calculate the estimate of HIV incidence ( $I_{est}$ ) using the sensitive/less

sensitive (S/LS, STARHS, and detuned) assay in screened populations [3].

In a hypothetical study of duration  $T_{obs}$  of an HIV testing population of size  $N$  who test at a mean intertest interval of  $T_{test}$  and have a true HIV incidence of  $I_{true}$  and is tested with a detuned assay with a window period of  $T_{win}$ , the Equation (1) can also be expressed by another equation replacing  $N_{dis}$  and  $N_{test}$  using only those variables ( $N$ ,  $T_{test}$ ,  $T_{obs}$ ,  $T_{win}$  and  $I_{true}$ ). This assumes that all temporal parameters ( $T$ ) are expressed in the same units and that the true incidence,  $I_{true}$ , and the mean intertest interval,  $T_{test}$ , is constant (uniform) across the study period ( $T_{obs}$ ) and that the duration of the study period is greater than or equal to the window period of the detuned assay ( $T_{obs} \geq T_{win}$ ).

In the absence of SCE:

From [3]

$$I_{est} = \frac{N_{dis}}{N_{test} T_{win}} \tag{1}$$

$$N_{test} = \frac{N T_{obs}}{T_{test}} \tag{2}$$

$$N_{dis} = \frac{N T_{obs} T_{win} I_{true}}{T_{test}} \tag{3}$$

(4) substitute (2) and (3) in (1)

$$I_{est} = \frac{\frac{N T_{obs} T_{win} I_{true}}{T_{test}}}{\frac{N T_{obs}}{T_{test}} T_{win}} = I_{true} \tag{4}$$

In the presence of SCE:

$N'_{test}$ , total number of tests in the period of observation in the presence of SCE;

$N'_{dis}$ , number of tests discordant [S(+)-LS(-)] in the presence of SCE and  $T_{win} \geq T_{sce}$ ;

$N''_{dis}$ , number of tests discordant [S(+)-LS(-)] in the presence of SCE and  $T_{win} < T_{sce}$ ;

$P_{sce}$ , proportion of newly infected persons having 'additional' testing in  $T_{sce}$  due to SCE;

$T_{sce}$ , time period used to define  $P_{sce}$

$$N'_{test} = \left( \frac{N T_{obs}}{T_{test}} \right) + ((N(1 - (1 - I_{true})^{T_{obs}})) P_{sce}) \tag{2'}$$

For  $T_{win} \geq T_{sce}$ ,

$$N'_{dis} = \left( \frac{N T_{obs} T_{win} I_{true}}{T_{test}} \right) + \left( \left( (N(1 - (1 - I_{true})^{T_{obs}})) - \left( \frac{N T_{obs} T_{win} I_{true}}{T_{test}} \right) \right) P_{sce} \right) \tag{3a'}$$

For  $T_{win} < T_{sce}$ ,

$$N''_{dis} = \left( \frac{N T_{obs} T_{win} I_{true}}{T_{test}} \right) + \left( \left( (N(1 - (1 - I_{true})^{T_{obs}})) - \left( \frac{N T_{obs} T_{win} I_{true}}{T_{test}} \right) \right) \left( \frac{P_{sce} T_{win}}{T_{sce}} \right) \right) \tag{3b'}$$

(4a') Substitute (2') and (3a') in (1)

For  $T_{win} \geq T_{sce}$ ,

$$I'_{est} = \frac{\left( \frac{N T_{obs} T_{win} I_{true}}{T_{test}} \right) + \left( \left( (N(1 - (1 - I_{true})^{T_{obs}})) - \left( \frac{N T_{obs} T_{win} I_{true}}{T_{test}} \right) \right) P_{sce} \right)}{\left( \left( \frac{N T_{obs}}{T_{test}} \right) + ((N(1 - (1 - I_{true})^{T_{obs}})) P_{sce}) \right) T_{win}} \tag{4a'}$$

(4b') Substitute (2') and (3b') in (1)

For  $T_{\text{win}} < T_{\text{sce}}$ ,

$$I''_{\text{est}} = \frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right) + \left(\left((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) \left(\frac{P_{\text{sce}} T_{\text{win}}}{T_{\text{sce}}}\right)\right)}{\left(\left(\frac{N T_{\text{obs}}}{T_{\text{test}}}\right) + ((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}}))P_{\text{sce}})\right) T_{\text{win}}} \quad (4b')$$

(5') Change in measured incidence due to SCE = (4a') minus (4)

$$I'_{\text{est}} - I_{\text{est}} = \frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right) + \left(\left((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) P_{\text{sce}}\right)}{\left(\left(\frac{N T_{\text{obs}}}{T_{\text{test}}}\right) + ((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}}))P_{\text{sce}})\right) T_{\text{win}}} - \left(\frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)}{\left(\frac{N T_{\text{obs}}}{T_{\text{test}}}\right) T_{\text{win}}}\right) \quad (5')$$

$$I'_{\text{est}} - I_{\text{est}} = \frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right) + \left(\left((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) P_{\text{sce}}\right)}{N'_{\text{test}} T_{\text{win}}} - \frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)}{N_{\text{test}} T_{\text{win}}}$$

If  $N'_{\text{test}} \approx N_{\text{test}}$  (i.e. there is a negligible number of additional tests due to SCE) then:

$$I'_{\text{est}} - I_{\text{est}} \approx \frac{\left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right) + \left(\left((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) P_{\text{sce}}\right) - \frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}}{N_{\text{test}} T_{\text{win}}}$$

$$I'_{\text{est}} - I_{\text{est}} \approx \frac{\left(\left((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) P_{\text{sce}}\right)}{N_{\text{test}} T_{\text{win}}}$$

$$I'_{\text{est}} - I_{\text{est}} \approx \frac{\left(\left((N(1 - (1 - I_{\text{true}})^{T_{\text{obs}}})) - \left(\frac{N T_{\text{obs}} T_{\text{win}} I_{\text{true}}}{T_{\text{test}}}\right)\right) P_{\text{sce}}\right)}{\frac{N T_{\text{obs}}}{T_{\text{test}}} T_{\text{win}}}$$

Therefore the change in the estimated incidence due to SCE is directly proportional to  $P_{\text{sce}}$  and the relationship between  $T_{\text{obs}}$ ,  $T_{\text{win}}$ ,  $T_{\text{test}}$  and  $I_{\text{true}}$  as indicated in (5').

## References

- McDougall L, Schechter MT, Sheps S, Craib KJP, Fay S, Montaner JSG. **Bias in AIDS cohort studies and an empiric demonstration of volunteer bias in an investigation of homosexual men** [abstract #WC95]. In: *7th International AIDS Conference*; Florence, Italy; 16–21 June 1991.
- Batter V, Matela B, Nsuami M, Manzila T, Kamenga M, Behets F, et al. **High HIV-1 incidence in young women masked by stable overall seroprevalence among childbearing women in Kinshasa, Zaire: estimating incidence from serial seroprevalence data.** *AIDS* 1994; **8**:811–817.
- Janssen RS, Satten GA, Stramer SL, Rawal BD, O'Brien TR, Weiblen BJ, et al. **New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes.** *JAMA* 1998; **280**:42–48.
- Brookmeyer R, Quinn T, Shepherd M, Mehendale S, Rodrigues J, Bollinger R. **The AIDS epidemic in India: a new method for estimating current human immunodeficiency virus (HIV) incidence rates.** *Am J Epidemiol* 1995; **142**:709–713.
- Pilcher CD, Fiscus SA, Nguyen TQ, Foust E, Wolf L, Williams D, et al. **Detection of acute infections during HIV testing in North Carolina.** *N Engl J Med* 2005; **352**:1873–1883.
- Schacker T, Collier AC, Hughes J, Shea T, Corey L. **Clinical and epidemiologic features of primary HIV infection.** *Ann Intern Med* 1996; **125**:257–264.
- Kahn JO, Walker BD. **Acute human immunodeficiency virus type 1 infection.** *N Engl J Med* 1998; **339**:33–39.
- Hecht FM, Busch MP, Rawal B, Webb M, Rosenberg E, Swanson M, et al. **Use of laboratory tests and clinical symptoms for identification of primary HIV infection.** *AIDS* 2002; **16**:1119–1129.
- Kinloch-de-Loes S, de Saussure P, Saurat JH, Stalder H, Hirschel B, Perrin LH. **Symptomatic primary infection due to human immunodeficiency virus type 1: review of 31 cases.** *Clin Infect Dis* 1993; **17**:59–65.
- Burchell AN, Calzavara L, Ramuscak N, Myers T, Major C, Rachlis A, et al. **Symptomatic primary HIV infection or risk experiences? Circumstances surrounding HIV testing and diagnosis among recent seroconverters.** *Int J STD AIDS* 2003; **14**:601–608.
- Stekler J, Collier AC, Holmes KK, Golden MR. **Primary HIV infection education: knowledge and attitudes of HIV-negative men who have sex with men attending a public health sexually transmitted disease clinic.** *J Acquir Immune Defic Syndr* 2006; **42**:123–126.
- Remis RS, Major C, Fearon M, Whittingham E, Degazio T. **Enhancing HIV diagnostic data for surveillance of HIV infection: use of the detuned assay** (abstract #MoPeC2340). In: *13th International AIDS Conference*; Durban, South Africa; 9–14 July 2000.
- Remis RS, Fikre M, Palmer RWH, Swantee C, Major C, Fearon M, et al. **Enhancing HIV diagnostic data for surveillance of HIV infection: results from the detuned assay to December 2002.** In: *5th Annual Research Conference*; Ontario HIV Treatment Network, Toronto, Canada; 4–5 November 2003.
- Remis RS, Major C, Calzavara L, Myers T, Burchell A, Whittingham EP. *The HIV epidemic among men who have sex with other men: the situation in Ontario in the year 2000.* Toronto: Department of Public Health Sciences, University of Toronto; 2000. 58 pp.
- Remis RS, Swantee C, Schiedel L, Liu J. *Report on HIV/AIDS in Ontario 2006.* Ontario Ministry of Health and Long-Term Care; 2008.
- Remis RS, Palmer RWH, Raboud J. **Estimates of HIV incidence based on detuned assay results may be strongly biased: evidence from a simulation study** (abstract #MoPeC3457). In: *14th International AIDS Conference*; Barcelona, Spain; 7–12 July 2002.
- Remis RS, Alary M, Otis J, Masse B, Demers E, Vincelette J, et al. **No increase in HIV incidence observed in a cohort of men who have sex with other men in Montreal.** *AIDS* 2002; **16**:1183–1185.
- Calzavara L, Burchell AN, Major C, Remis RS, Corey P, Myers T, et al. **Increases in HIV incidence among MSM undergoing repeat diagnostic testing in Ontario, Canada.** *AIDS* 2002; **16**:1655–1661.
- Hogg RS, Weber AE, Chan K, Martindale S, Cook D, Miller ML, Craib KJ. **Increasing incidence of HIV infections among young gay and bisexual men in Vancouver.** *AIDS* 2001; **15**:1321–1322.
- Holmberg SD. **The estimated prevalence and incidence of HIV in 96 large US metropolitan areas.** *Am J Public Health* 1996; **86**:642–654.
- Kothe D, Byers RH, Caudill SP, Satten GA, Janssen RS, Hannon WH, Mei JV. **Performance characteristics of a new less sensitive HIV-1 enzyme immunoassay for use in estimating HIV seroincidence.** *J Acquir Immune Defic Syndr* 2003; **33**:625–634.
- Remis RS, Major C, Swantee C, Fearon M, Wallace E, Whittingham E. **Laboratory Enhancement Study: using the detuned assay to determine HIV incidence in Ontario – results and methodologic perspectives.** In: *Workshop on STARHS Assay*; U.S. Centers for Disease Control and Prevention, Albany, New York, USA; 14–15 November 2001.
- Norton J, Elford J, Sherr L, Miller R, Johnson MA. **Repeat HIV testers at a London same-day testing clinic.** *AIDS* 1997; **11**:773–781.
- Weinstock H, Dale M, Gwinn M, Satten GA, Kothe D, Mei J, et al. **HIV seroincidence among patients at clinics for sexually transmitted diseases in nine cities in the United States.** *J Acquir Immune Defic Syndr* 2002; **29**:478–483.
- Murphy G, Parry JV, Gupta SB, Graham C, Jordan LF, Nicoll AN, Gill ON. **Test of HIV incidence shows continuing HIV transmission in homosexual/bisexual men in England and Wales.** *Commun Dis Public Health* 2001; **4**:33–37.
- Schwarzc S, Kellogg T, McFarland W, Louie B, Kohn R, Busch M, et al. **Differences in the temporal trends of HIV seroincidence and seroprevalence among sexually transmitted disease clinic patients, 1989–1998: application of the serologic testing algorithm for recent HIV seroconversion.** *Am J Epidemiol* 2001; **153**:925–934.
- Murphy G, Parry JV, Jordan LF, Charlet A, Gill ON. **STARHS test shows HIV transmission continuing in UK despite widespread HAART** (abstract #TuPeC4868). In: *14th International AIDS Conference*; Barcelona, Spain; 7–12 July 2002.
- Schwarzc SK, Kellogg TA, McFarland W, Louie B, Klausner J, Withum DG, Katz MH. **Characterization of sexually transmitted disease clinic patients with recent human immunodeficiency virus infection.** *J Infect Dis* 2002; **186**:1019–1022.
- Withum DG, McFarland W, Kellogg T, Dilley J, Adler B, Sabatino J, et al. **Acceptance of a test to detect early HIV-1 infection (STARHS) and calculation of HIV-1 incidence among persons attending anonymous test sites** (abstract #MoPeB2184). In: *13th International AIDS Conference*; Durban, South Africa; 9–14 July 2000.
- Kellogg T, McFarland W, Dilley J, Rinaldi J, Adler B, Sabatino J, et al. **Three methods to measure HIV incidence among persons seeking anonymous testing** (abstract #MoPeC2436). In: *13th International AIDS Conference*; Durban, South Africa; 9–14 July 2000.
- McFarland W, Busch MP, Kellogg TA, Rawal BD, Satten GA, Katz MH, et al. **Detection of early HIV infection and estimation of incidence using a sensitive/less-sensitive enzyme immunoassay testing strategy at anonymous counseling and testing sites in San Francisco.** *J Acquir Immune Defic Syndr* 1999; **22**:484–489.
- Kral AH, Lorvick J, Gee L, Bacchetti P, Rawal B, Busch M, Edlin BR. **Trends in human immunodeficiency virus seroincidence among street-recruited injection drug users in San Francisco, 1987–1998.** *Am J Epidemiol* 2003; **157**:915–922.
- Martindale SL, Cook D, Weber AE, Miller ML, Chan K, Craib KJP, Hogg RS. **The impact of STARHS 'detuned assay' results on HIV incidence calculations in an ongoing cohort of men who have sex with men (MSM) in Vancouver** [abstract #Tu-PeC4870]. In: *14th International AIDS Conference*; Barcelona, Spain; 7–12 July 2002.
- Busch MP, Rawal BD, Stramer SL, Aberle-Grasse J, Schreiber GB, Janssen RS. **Estimation of HIV incidence in US blood donors using a novel detuned anti-HIV EIA test strategy** [abstract #531]. In: *5th Conference on Retroviruses and Opportunistic Infections*; Chicago, Illinois, USA; 1–5 February 1998.

35. Busch MP, Aberle-Grasse J, Rawal BD, Watanabe K, Schreiber GB, Satten GA, Janssen RS. **Demographic correlates of HIV incidence among US blood donors** (abstract #272). In: *6th Conference on Retroviruses and Opportunistic Infections*; Chicago, USA; 31 January–4 February 1999.
36. Glynn SA, Kleinman SH, Schreiber GB, Busch MP, Wright DJ, Smith JW, et al. **Trends in incidence and prevalence of major transfusion-transmissible viral infections in US blood donors, 1991 to 1996. REDS Study.** *JAMA* 2000; **284**:229–235.
37. Busch MP, Glynn SA, Stramer SL, Strong DM, Caglioti S, Wright DJ, et al. **A new strategy for estimating risks of transfusion-transmitted viral infections based on rates of detection of recently infected donors.** *Transfusion* 2005; **45**:254–264.
38. Hall HI, Song R, Rhodes P, Prejean J, An Q, Lee LM, et al. **Estimation of HIV incidence in the United States.** *JAMA* 2008; **300**:520–529.
39. Karon JM, Song R, Brookmeyer R, Kaplan EH, Hall HI. **Estimating HIV incidence in the United States from HIV/AIDS surveillance data and biomarker HIV test results.** *Stat Med* 2008; **27**:4617–4633.
40. Lee SH, Sheon N. **Responsibility and risk: accounts of reasons for seeking an HIV test.** *Social Health Illn* 2008; **30**: 167–181.