Published in final edited form as:
Sex Transm Infect. 2010 August ; 86(4): 254–257. doi:10.1136/sti.2009.037291.

Interpopulation variation in HIV testing promptness may introduce bias in HIV incidence estimates using the serologic testing algorithm for recent HIV seroconversion

Edward White[1], **Gary Goldbaum**[2], **Steven Goodreau**[3], **Thomas Lumley**[4], and **Stephen E Hawes**[2]

[1]Yale University, School of Public Health, Center for Interdisciplinary Research on AIDS, New Haven, Connecticut, USA

[2]Department of Epidemiology, University of Washington, Seattle, Washington, USA

[3]Department of Anthropology, University of Washington, Seattle, Washington, USA

[4]Department of Biostatistics, University of Washington, Seattle, Washington, USA

Abstract

Objectives—The serologic testing algorithm for recent HIV seroconversion (STARHS) calculates incidence using the proportion of testers who produce a level of HIV antibody high enough to be detected by ELISA but low enough to suggest recent infection. The validity of STARHS relies on independence between dates of HIV infection and dates of antibody testing. When subjects choose the time of their own test, testing may be motivated by risky behaviour or symptoms of infection and the criterion may not be met. This analysis was conducted to ascertain whether estimates of incidence derived using STARHS were consistent with estimates derived using a method more robust against motivated testing.

Methods—A cohort-based incidence estimator and two STARHS methods were applied to identical populations (n=3821) tested for HIV antibody at publicly funded sites in Seattle. Overall seroincidence estimates, demographically stratified estimates and incidence rate ratios were compared across methods. The proportion of low-antibody testers among HIV-infected individuals was compared with the proportion expected given their testing histories.

Results—STARHS estimates generally exceeded cohort-based estimates. Incidence ratios derived using STARHS between demographic strata were not consistent across methods. The proportion of HIV-infected individuals with lower antibody levels exceeded that which would be expected under independence between infection and testing.

Conclusions—Incidence estimates and incidence rate ratios derived using methods that rely on the changing antibody level over the course of HIV infection may be vulnerable to bias when applied to populations who choose the time of their own testing.

Correspondence to :Dr Edward White, Yale University School of Public Health, Center for Interdisciplinary Research on AIDS, 60 College Street, PO Box 208034, New Haven, CT 065520-8034, USA; e.white@yale.edu.

Contributors EW was the analyst and writer for this manuscript and used it as a major part of his dissertation in epidemiology at the University of Washington. GG was the principal investigator of the project from which the data for this analysis were obtained and provided assistance with the manuscript. SG provided guidance in biostatistical methods. TL also provided guidance in biostatistical methods. SH provided assistance with the manuscript.

Competing interests None.

Ethics approval The University of Washington Human Subjects Division approved this analysis. The Washington State Human Research Review Section and the CDC Human Research Protection Office approved the parent study.

An important epidemiological measure of HIV transmission is seroincidence.[1] HIV seroincidence estimation methods are few and are limited by bias and lack of generalisability. Seroincidence has been estimated from cohorts of initially negative, repeat testers; this method is expensive due to relatively low incidence and is subject to retention bias.[2–4] Back-calculation methods are vulnerable to bias through variation in the diagnosis of infection and in the distribution of times from HIV acquisition to the development of symptoms. Furthermore, the introduction of highly active antiretroviral therapy, has greatly limited the reliability of back-calculation methods.[5,6] In 1998, Janssen et al[7] reported the development of a method incorporating a new test, applied to single specimens from HIV-infected testers, which could be used to estimate seroincidence without the expense and bias associated with other incidence estimation methods.

That method, the serologic testing algorithm to detect recent HIV seroconversion (STARHS), uses a weakened version of the ELISA, and either the less sensitive enzyme immunoassay (LS-EIA) or the BED. To develop STARHS, the LS-EIA was applied to stored diagnostic specimens from persons with known negative ELISA test dates and subsequent positive ELISA test dates. The developers used the standardised optical densities (SOD) from those results, and the test dates associated with them, to model two values: the SOD, and the corresponding estimated mean number of days following the production of antibodies detectable by a standard ELISA at which the an LS-EIA result would equal or surpass that SOD, which together would minimise misclassification. The mean number of days between production of antibody detectable by ELISA and sufficient for the cut-off SOD was termed $\bar{\omega}$ and was estimated to have a mean of 129 days (95% CI 109 to 149 days). ELISA-positive testers whose SOD is below the cut-off on LS-EIA are termed ‘LS-EIA non-reactive’ (see supplementary figure 1, available online only).

STARHS was devised to estimate annual incidence by multiplying the proportion of LS-EIA non-reactive results among testers by $365/\bar{\omega}$. In order to estimate the proportion of testers who became infected in a given period, STARHS assumes independence between the time of HIV infection and testing. This assumption may not be met in clinical settings. Some individuals who test HIV antibody positive do so following symptoms of early HIV infection or with the suspicion that he or she has recently become infected.[8,9] In addition, testing may be prompted by symptoms of a bacterial sexually transmitted infection near the time of HIV acquisition.[10]

This analysis applies STARHS to men who have sex with men (MSM) who sought HIV testing at Public Health—Seattle and King County (PHSKC). The authors sought to determine whether, using identical datasets from a clinical setting in which individuals chose the time of their testing, STARHS-based incidence estimates were similar to cohort-based incidence estimates. They also tested whether incidence ratios taken from data stratified by age and race/ethnicity, would be different when using STARHS and the cohort method. To test for possible lack of independence between HIV acquisition and testing, the authors also compared the proportion of LS-EIA non-reactive testers among all ELISA-positive testers with the proportion that would be expected given the distribution of intervals between positives’ last negative ELISA result and the first positive ELISA result.

Methods

Setting

PHSKC provides over 10 000 HIV tests annually at low or no cost through over 20 locations and through outreach activities aimed at high-prevalence populations (eg, MSM and drug injectors). The proportion of tests performed on MSM varies year to year, but is generally one in four. Nearly two-in-five HIV tests are performed on patients of the PHSKC Sexually

Transmitted Disease Clinic; however, among MSM that proportion is one in four. Whereas the positive test prevalence for all testers is 1.2%, MSM positive test prevalence is 4.4%. Testing records are linkable by individual within the PHSKC system; anonymous testers are assigned a unique code based on easily recalled characteristics: mother's maiden name, city of birth, etc; over time this system has proved highly reliable. Seroincidence estimates using STARHS and the cohort method were. The University of Washington Human Subjects Division approved this analysis. The Washington State Human Research Review Section and the CDC Human Research Protection Office approved the parent study.

Laboratory

Laboratory methods for the Abbott LS-3A11 have been described in detail elsewhere.[7] Briefly, blood specimens testing positive by repeat ELISA and confirmatory western blot are retested by the Abbott 311A LS-EIA. (For brevity, specimens testing positive by ELISA and confirmed by western blot will hereafter be referred to as 'ELISA/WB-positive'.) As per the Centers for Disease Control and Prevention (CDC) protocol, specimens with LS-EIA non-reactive results were retested with the assay. Blood specimens testing positive by ELISA and confirmed by western blot, collected through public HIV testing sites in Seattle between 1 January 1996 and 13 November 2000, were placed in storage at −40°C, except for a 4-month period in which collection was suspended. The PHSKC laboratory followed CDC protocol for this study, testing blinded, stored specimens.[11]

Testing records data

STARHS estimates incidence for the period in which stored ELISA/WB-positive specimens later tested by LS-EIA were collected, and uses test records only from that period. The cohort-based method uses records that precede or follow the incidence estimation period. Only test records that met the inclusion criteria for use in both the cohort method and STARHS methods, and for the same period, were used. In sum, records used in this analysis were characterised by the following: (1) the individual was tested at least once for the period for which seroincidence was estimated; (2) had at least two test records in the PHSKC database; (3) had not previously tested positive; (4) was MSM and (5) was 18 years or older. A fuller discussion of the inclusion of test records used in this analysis is given in supplemental materials 1 (available online only).

Analyses

Seroincidence estimates using three methods were compared, for the same incidence estimation period, 1 January 1996 to 13 November 2000. Incidence estimates stratified by age category and by race/ethnicity were also calculated. The cohort method that served as a referent standard estimate was adapted from a technique developed by Kitayaporn et al.[12] Person-time contributed by each tester i began with his earliest known ELISA-negative test or the start of the incidence estimation period if that test occurred earlier. For testers with no ELISA-positive result recorded, person-time concluded with the last ELISA-negative test or the end of the incidence estimation period if that test was later. For individuals who tested ELISA/WB-positive, the period between the last ELISA-negative result and the first ELISA/WB-positive result was considered a seroconversion interval. The number of cases occurring in an observation period was the sum of the proportions of each seroconversion period falling within the observation period. The amount of person-time in which these cases occurred was estimated as the sum of time between ELISA-negative results occurring within the incidence estimation period, plus the expected time between the last ELISA-negative result and the production of antibodies detectable by ELISA. Estimates of seroincidence were calculated using STARHS, and using Satten's correction for the overcontribution of a LS-EIA non-reactive result from those who tested following a short interval after their most recent ELISA-negative result.[13] A fuller description of all three methods of seroincidence

estimation is given in supplemental materials 2 (available online only). We calculated incidence ratios between age strata and racial ethnic strata using all three methods. Finally, to assess independence between acquisition of infection and testing we calculated the expected proportion of LS-EIA non-reactive results based on the distribution of intervals between the records of the last ELISA-negative result and the first ELISA/WB-positive result. The expected number of LS-EIA non-reactive results was calculated as $\bar{\omega}$ days multiplied by the number of valid LS-EIA results, divided by the sum of days in the aforementioned intervals. A detailed description of the method is given in supplemental materials 2 (available online only).

Results

Of 3821 initially negative, repeat testers, 79 (2.1%) tested ELISA/WB-positive during the seroincidence estimation period. Six stored specimens produced invalid LS-EIA results or were of quantity insufficient for testing. Valid LS-EIA results were available for 73 individuals. Incidence estimates are presented in table 1. For all testers during the overall period for which incidence was estimated, the ratio of point estimates of incidence derived using the crude STARHS method to those derived using the cohort-based method was 1.5.

For all age and race/ethnicity strata except Latino ethnicity, point estimates for crude STARHS incidence estimates appreciably exceeded cohort incidence estimates. However, in every case, confidence intervals overlapped. STARHS incidence estimates obtained using Satten's correction for frequent testing were less than 10% greater than those obtained using crude STARHS, but further from the cohort estimator referent. Under STARHS, incidence ratios between age and racial/ethnic strata differed appreciably from incidence ratios calculated using the cohort method. Incidence ratios comparing younger MSM with older MSM were generally greater under STARHS methods. Under both STARHS methods and the cohort method, incidence was greater among Latino MSM than white MSM, but under both STARHS methods the incidence ratio between Latino MSM and white MSM was 2.5 times greater than under the cohort method.

Forty-one specimens of 73 with valid LS-EIA results were non-reactive, an observed proportion of 0.56 (95% CI 0.45 to 0.68). The expected proportion of LS-EIA non-reactive results based upon the interval between last ELISA-negative and first ELISA/WB-positive results was 0.42 (95% CI 0.30 to 0.53) for an observed-to-expected ratio of 1.33. Ad-hoc analyses were conducted limiting observations to individuals for whom the interval between the last ELISA-negative result and the first ELISA/WB-positive result was greater than $\bar{\omega}$; for this subset the observed proportion of LS-EIA non-reactive results was 0.54 (95% CI 0.42 to 0.65) and the expected proportion was 0.38 (95% CI 0.27 to 0.50) for an observed-to-expected ratio of 1.42. In addition, we limited analysis to the 63% of testers for whom the seroconversion interval was greater than 2 $\bar{\omega}$; for this subset the observed proportion of LS-EIA non-reactive results was 0.50 (95% CI 0.36 to 0.64) and the expected proportion was 0.25 (95% CI 0.12 to 0.37) for an observed-to-expected ratio of 2.00. Observations were then limited to the 51 individuals testing ELISA/WB positive at the two largest-volume PHSKC testing sites, which routinely gather self-reported data on clients' date of last HIV test and result. In this subanalysis, the observed proportion of LS-EIA non-reactive results was 0.59 (95% CI 0.45 to 0.73) and the expected proportion was 0.38 (0.24 to 0.51) for a ratio of 1.55.

Discussion

Seroincidence estimates derived from STARHS exceeded estimates from the cohort-based method by more than 50%. Although more precise than crude STARHS incidence estimates,

corrected STARHS incidence estimates yielded even higher incidence estimates. Upon stratification by age and race/ethnicity, incidence ratios of STARHS-based estimates were not consistent with incidence ratios of cohort-based estimates. Because for some subgroups testing and infection may be more independent and for others it may be less, or may vary on other factors that affect the likelihood of testing during the period one would test LS-EIA non-reactive, stratification may introduce confounding. Finally, the proportion of LS-EIA non-reactive results among ELISA/WB-positive testers was greater than would be expected under independence between infection and testing.

This comparison of STARHS and cohort-based seroincidence estimates has a high degree of validity. Seroincidence was calculated for both methods using identical sets of individuals whose records met the inclusion criteria for both methods. Seroincidence was calculated for identical periods, in identical clinical settings. The cohort estimator, measuring events under observation, is a sound referent against STARHS and is robust against bias from motivated testing, as reducing ELISA/WB-positive testers' seroconversion periods would be small compared with the person-time contributed by other testers (E White, T Lumley, S Goodreau, et al, 2009, unpublished data).

These analyses comparisons were subject to a number of limitations. Had it been possible, all testers who presented for reasons motivated by potential exposure would have been excluded from analysis, including those presenting with sexually transmitted disease symptoms or as a result of partner notification. However, data indicating the reason for presentation were not available. In addition, HIV incidence is low, even in high-prevalence populations, providing little power for stratified analyses of data from a single city.

This analysis compares STARHS-based incidence estimates with those from a cohort estimator that measured person-time and disease events under observation, and in identical sets of repeat testers in which subjects chose the timing of testing. The approach differs from other comparisons of STARHS-based and cohort-based estimates of incidence, which used testers' self-report of a previous negative test, many of which occurred at a location for which records of all previous negative tests were unavailable and outside the study frame.[14,15] Had those previous tests at locations for which records were unavailable had positive results, they would not have been counted as disease events; this circumstance violates the concept of the study base.

Song et al,[16] in a previous analysis of a database overlapping that used in this analysis, found the observed proportion of LS-EIA non-reactive results to agree with the expected proportion. However, that analysis included MSM who tested ELISA/WB-positive at PHSKC, but may not have had a previous test through PHSKC, and included only those testing positive at the two largest-volume locations in the PHSKC system, both of which routinely collected the reported date of client's last HIV antibody test and the result of that test. In addition, that analysis used an unpublished estimate of $\bar{\omega}$ of 140 days. However, the ad-hoc analysis limited to those testing positive at the same locations yielded a greater proportion of observed LS-EIA non-reactive results than would be expected. It is possible that the degree of motivated testing among MSM who consistently test through PHSKC is greater than for those who do not.

Karon and colleagues[17] recently proposed a correction for the detection of HIV infection within the window period of a biomarker test suggesting recent infection. The correction assesses the probability of testing very soon after acquiring HIV. However, members of all strata of interest may not test in response to the suspicion of infections, of symptoms, or isolated risk in the same manner. Some may test sooner after the acquisition of infection than others. McDougal et al [18] have suggested an additional correction to control for LS-

EIA non-reactive results for a some persons with long-standing HIV infection and reduced immune function. The capacity of these adjustments to overcome misclassification bias is controversial. [19] Therefore, it will be important that any correction to attenuate these biases apply to all strata. It will be of interest to apply their correction to the current dataset to determine whether controlling for the period between the last ELISA-negative result and the first ELISA/WB-positive result is sufficient to address the large ratio of observed-to-expected LS-EIA non-reactive results.

To address the persistent transmission of HIV, seroincidence among populations and subpopulations must be quantified in order to identify priorities of need. STARHS may offer an opportunity to produce more generalisable estimates of seroincidence by requiring only one specimen from subjects. However, in the present analysis, incidence ratios between groups of interest varied by the method of estimation. Because surveillance data partly determine prevention resource allocation, we must view sound incidence estimation methodologies as important as evidence-based prevention techniques.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank Drs John Karon, Riquang Song and Robert Byers, of (or formerly of) the Centers for Disease Control and Prevention for their kind advice in the preparation of this manuscript and statistical advice. They also express gratitude to Dr Matthew Golden of the Seattle STD Clinic for his assistance.

Funding The parent study providing the data used in this analysis was conducted with support of the Centers for Disease Control and Prevention, grant number U62/CCU006260. This analysis was conducted with the support of the University of Washington, Department of Epidemiology and by grant number T32 MH020031 from the National Institute of Mental Health. Content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

References

1. Weinstock H, Sweeney S, Satten GA, et al. HIV incidence and risk factors among patients repeatedly tested for HIV attending sexually transmitted disease clinics in the United States, 1991 to 1996. J Acquir Immune Defic Syndr Hum Retro. 1998; 19:506–12.
2. Catania JA, Coates TJ, Stall R, et al. Changes in condom use among homosexual men in San Francisco. Health Psychol. 1991; 10:190–9. [PubMed: 1879391]
3. McFarland W, Kellogg TA, Dilley J, et al. Estimation of human immunodeficiency virus (HIV) incidence among repeat anonymous testers in San Francisco. Am J Epidemiol. 1997; 146:662–4. [PubMed: 9345120]
4. Karon JM, Fleming PL, Steketee RW, et al. HIV in the United States at the turn of the century: an epidemic in transition. Am J Public Health. 2001; 91:1060–8. [PubMed: 11441732]
5. Karon JM, Khare M, Rosenberg PS. The current status of methods for estimating the prevalence of human immunodeficiency virus in the United States of America. Stat Med. 1998; 17:127–42. [PubMed: 9483724]
6. Centers for Disease ControlPrevention (CDC). HIV prevalence estimates–United States, 2006. MMWR Morb Mortal Wkly Rep. 2008; 57:1073–6. [PubMed: 18830210]
7. Janssen RS, Satten GA, Stramer SL, et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. JAMA. 1998; 280:42–8. [PubMed: 9660362]
8. Burchell AN, Calzavara L, Ramuscak N, et al. Symptomatic primary HIV infection or risk experiences? Circumstances surrounding HIV testing and diagnosis among recent seroconverters. Int J STD AIDS. 2003; 14:601–8. [PubMed: 14511496]

9. Wenger NS, Kusseling FS, Beck K, et al. When patients first suspect and find out they are infected with the human immunodeficiency virus: implications for prevention. AIDS Care. 1994; 6:399–405. [PubMed: 7833357]
10. Remis RS, Palmer RW. Testing bias in calculating HIV incidence from the serologic testing algorithm for recent HIV seroconversion. AIDS. 2009; 23:493–503. [PubMed: 19240458]
11. Centers for Disease Control and Prevention. Protocol: human immunodeficiency virus type 1 low sensitivity enzyme immunoassay. Atlanta: Modified Abbott; 1998. p. 3A11
12. Kitayaporn D, Uneklabh C, Weniger BG, et al. HIV-1 incidence determined retrospectively among drug users in Bangkok, Thailand. AIDS. 1994; 8:1443–50. [PubMed: 7818815]
13. Satten GA, Janssen R, Busch MP, et al. Validating marker-based incidence estimates in repeatedly screened populations. Biometrics. 1999; 55:1224–7. [PubMed: 11315072]
14. Kellogg TA, Loeb L, Dilley J, et al. Comparison of three methods to measure HIV incidence among persons seeking voluntary, anonymous counseling and testing. J Acquir Immune Defic Syndr. 2005; 39:112–20. [PubMed: 15851921]
15. Nash D, Bennani Y, Ramaswamy C, et al. Estimates of HIV incidence among persons testing for HIV using the sensitive/less sensitive enzyme immunoassay. New York City, 2001. J Acquir Immune Defic Syndr. 2005; 39:102–11. [PubMed: 15851920]
16. Song R, Karon JM, White E, et al. Estimating the distribution of a renewal process from times at which events from an independent process are detected. Biometrics. 2006; 62:834–46.
17. Karon JM, Song R, Brookmeyer R, et al. Estimating HIV incidence in the United States from HIV/AIDS surveillance data and biomarker HIV test results. Stat Med. 2008; 27:4617–33. [PubMed: 18833636]
18. 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 Retrovirol. 2006; 22:945–52.
19. Brookmeyer R. Should biomarker estimates of HIV incidence be adjusted? AIDS. 2009; 23:485–91. [PubMed: 19165087]

Key messages

Seroincidence estimates using STARHS are vulnerable to bias when applied to specimens and data from individuals who chose the time of their testing.

When applied to identical sets of testers and incidence estimation periods, STARHS seroincidence estimates are generally higher than cohort-based seroincidence estimates.

Bias arises from infected individuals testing sooner after HIV acquisition than would be expected under independence between acquisition and testing, violating an assumption of STARHS.

Seroincidence estimates stratified by age and race/ethnicity may be confounded by variation in the interval between HIV acquisition and testing, between strata.

Table 1

HIV seroincidence estimates by estimation method, stratified by age and race/ethnicity: new infections per 100 person-years, 95% CI and ratios between stratified incidence estimates

	Cohort method		Crude STARHS		Corrected STARHS*	
Population	**Incidence (95% CI)**	**IRR**	**Incidence (95% CI)**	**IRR**	**Incidence (95% CI)**	**IRR**
Total						
(n=3821)	1.14 (0.89 to 1.43)		1.70 (0.98 to 2.51)		176 (1.04 to 2.75)	
Age group, years						
(n=3447)						
20–29	1.39 (0.95 to 2.04)	2.96	1.84 (0.81 to 3.77)	4.18	1.91 (0.87 to 3.66)	4.24
30–39	1.15 (0.80 to 1.63)	2.44	2.38 (1.21 to 3.89)	5.41	2.45 (1.30 to 4.22)	5.44
40–49	0.47 (0.16 to 0.97)	Ref	0.44 (0.04 to 1.71)	Ref	0.45 (0.05 to 1.86)	Ref
Race						
(n=3516)						
White	1.05 (0.79 to 1.37)	Ref	1.75 (0.97 to 2.67)	Ref	1.81 (1.05 to 2.90)	Ref
Black	1.30 (0.43 to 3.07)	1.24	2.92 (0.64 to 8.05)	1.67	3.04 (0.83 to 8.85)	1.68
Latino	3.24 (1.29 to 6.62)	3.09	2.17 (0.36 to 6.84)	1.24	2.26 (0.47 to 7.49)	1.25

*Crude serological testing algorithm to detect recent HIV seroconversion (STARHS) and corrected STARHS adjusted for six missing or unsatisfactory less sensitive enzyme immunoassay (LS-EIA) results. The formula for corrected STARHS is given by Satten *et al.*[13]

IRR, incidence rate ratio.