

IBS

POPULATION PROGRAM ■
INSTITUTE OF BEHAVIORAL SCIENCE ■
UNIVERSITY OF COLORADO AT BOULDER ■



WORKING PAPER

Disclosure of HIV Test Results and Non-response Bias in Seroprevalence Surveys

Georges Reniers
Tekebash Araya
Yemane Berhane
Gail Davey
Eduard J. Sanders

August 2007

Disclosure of HIV Test Results and Non-response Bias in Seroprevalence Surveys

Georges Reniers^{1 2 3}
Tekebash Araya^{4 5}
Yemane Berhane⁴
Gail Davey⁴
Eduard J. Sanders^{6 7}

August 2007

Acknowledgements: The study was funded by the AIDS Foundation of Amsterdam (grant 7022) and received support from the Centralized School of Nursing and the Department of Community Health of Addis Ababa University, and the Health Service Amsterdam (GGD Amsterdam). The study was in large part conducted while Georges Reniers was a PhD Student at the Population Studies Center and the Department of Sociology of the University of Pennsylvania. Georges Reniers also received support through a Hewlett Foundation grant to the University of Colorado at Boulder for the African Population Studies Research and Training Program. We like to thank the direction of the Zewditu Memorial Hospital for facilitating the surveillance, and the nurses and VCT team for their dedicated work. Admission diagnoses were coded with the assistance of Dr. Abiy Arefeayine and Nurse Misganaw Getaw. Tarik Tadesse and Yeshi G/Wold oversaw data collection and data entry. We benefited from thoughtful comments and methodological advice from Jimi Adams, Doug Ewbank, Richard Rogers, Rania Tfaily, the late Etienne van de Walle, and Susan Watkins.

Corresponding author:

Georges Reniers; Population Program; Institute of Behavioral Science; University of Colorado; 1424 Broadway, IBS #3; Boulder, CO 80309-0484; e-mail: reniers@colorado.edu

¹ University of Colorado at Boulder, Institute of Behavioral Science, Population Program

² University of the Witwatersrand, School of Social Sciences

³ University of the Witwatersrand, School of Public Health

⁴ Addis Ababa University, Department of Community Health

⁵ Addis Ababa University, Centralized School of Nursing

⁶ Kenya Medical Research Institute, Centre for Geographic Medicine Research

⁷ University of Oxford, Centre for Clinical Vaccinology and Tropical Medicine

Abstract

Objective(s): Investigate the effect of different study protocols with respect to the disclosure of HIV test results on non-response bias in HIV prevalence estimates.

Design: Nine-month surveillance of hospital admissions in Addis Ababa in which patients were approached for an HIV test. Patients had the choice between three consent levels: testing and post-test counseling (level A); testing without post-test counseling (level B); and total refusal (level C). For all patients, information was collected on basic sociodemographic background characteristics and admission diagnosis. That information is used to predict HIV status in those who refuse testing.

Methods: We first investigate the covariates of different levels of consent. We then quantify bias in HIV prevalence estimates due to refusal for testing via Heckman regression models that account for sample selection.

Results: We find that refusal positively correlates with the likelihood of infection and that non-response bias in HIV prevalence surveys depends on the study protocol: if disclosure of HIV status is implied in study participation, the bias is likely to be much larger than in a scenario where respondents can opt out of post-test counseling. We also find that consent for testing increased since the introduction of antiretroviral therapy in Ethiopia. Other covariates of refusal are age, gender, marital status, educational status and counselor.

Conclusions: Disclosure or non-disclosure of test results is an important consideration in studies that wish to minimize non-response bias in HIV prevalence surveys. The availability of ART is likely to reduce refusal rates.

Keywords

HIV prevalence estimates, non-response, selection bias, voluntary counseling and testing (VCT), Ethiopia, Addis Ababa, sub-Saharan Africa

1. Background and objectives

Expanding resources and progress in medical technology has brought HIV testing increasingly within reach of nationally representative surveys. This has generated new prospects for resolving bias in HIV prevalence estimates that are based on antenatal clinic (ANC) sentinel surveillance data, or, to provide a new gold standard for HIV prevalence estimates altogether (1-5). Data from population-based surveys are indeed valuable additions to ANC derived estimates, but they are also subject to bias due to limitations of the sampling frame (e.g., the exclusion of high risk groups in army barracks, prisons or migrant worker hostels) and non-response because of population mobility and refusal. The association of population mobility with HIV infection has been documented extensively (6-14). In comparison, relatively little is known about the relationship between refusal and HIV infection in nationally representative surveys that include a testing component (2, 3, 5). Several small-scale studies in STD and antenatal clinics concluded that refusals are positively associated with HIV status (15-23). Three studies remain inconclusive about the nature of the relationship or even suggest the opposite pattern (24-26). On aggregate, population-based surveys are believed to underestimate true HIV prevalence, but the few studies that addressed this issue have failed to identify significant bias due to refusal (5, 27-29).

A study design feature that often distinguishes health facility-based studies from population-based surveys and that may contribute to differences in non-response bias is the protocol for post-test counseling and disclosure of test results. Post-test counseling is often implied in study participation in health facility-based studies. This contrasts with most population-based surveys (e.g., the Demographic and Health Surveys) that commonly follow a protocol whereby respondents or clients are not informed about their HIV status. Instead, they are given a voucher to visit the nearest VCT center at no cost (if the service is not already free of charge) (30). In early studies involving testing for HIV, the disclosure of test results was not usually an option because samples had to be shipped to an off-site lab for analysis. With the increasing availability and reliability of rapid tests, post-test counseling is feasible in the same session in which the specimens are collected. Because of the ethical prescription that study participants should share in the benefits of research (30), the pressure to provide post-test counseling in nationally representative HIV prevalence surveys is likely to increase in the near future. Therefore it is important to assess how best to accommodate this prescription in the study protocol while maintaining the external validity of the ensuing HIV prevalence estimates.

In this study we evaluate the effect of different protocols by allowing for three 'consent levels': A) consent to testing and post-test counseling, B) consent to testing without disclosure of results and C) total refusal. We first study the covariates of the three consent levels. Via regression models that account for

sample selection, we then quantify the magnitude of bias in HIV prevalence estimates under different study protocols. The data come from surveillance in a governmental hospital in Addis Ababa. An advantage of a hospital sample over a random population sample is that it provides greater detail on the medical condition of respondents: whether they agree to testing or not. Because health status is a good predictor of HIV status, it can be used to assess the association between refusal and HIV status. In community-based studies, most measured traits correlate weakly with HIV status and this renders assessments of bias in prevalence estimates less reliable. The downside of our sample is that medical facility-based studies are not necessarily representative for the determinants of participation in population-based surveys. A final noteworthy feature of our study is that antiretroviral therapy (ART) was introduced in the hospital under surveillance during the course of data collection allowing us to evaluate its impact on willingness to be tested.

2. Setting

As many other urban centers in East Africa, Addis Ababa is severely affected by HIV/AIDS. Reported prevalence rates do, however, vary considerably by source: for 2003, the Ministry of Health estimated HIV prevalence at 14.6 percent (31). The 2005 Ethiopia Demographic and Health Survey (DHS) established HIV prevalence at 5.0 percent (32). The implications for the health infrastructure are nonetheless serious: in 2001, around half of all adult hospital deaths (age 13+) were attributed to the AIDS-related complex (33).

In July 2003, the Ethiopian government adopted a policy for the provision of ART, first via a co-pay scheme, later for free. By February 2006 close to 27,000 patients had ever been enrolled in an ART program, half of them in the capital (34).

3. Data and methods

Surveillance of hospital admissions and outpatient visits was initiated in the Zewditu Memorial Hospital in May 2003 and continued for nine months. Zewditu is a government medical facility in the inner city and was one of the few hospitals with a voluntary counseling and testing (VCT) center of sufficient capacity to accommodate our study. Initially, the surveillance covered the TB-HIV clinic (TB, ambulatory patients only), the medical emergency (ER), internal medicine (IM), gynaecology (GY) and pediatric wards (PE). For each patient, a ward nurse collected basic background characteristics as well as the admission and discharge diagnosis. One month into the study, the surgical ward (SU) was included in the surveillance and we added educational status, birthplace and marital status as additional information to be collected for each

patient. After new patients were identified, a ward nurse contacted the coordinator of the VCT unit who assigned a VCT-nurse to conduct pre-test counseling and ask for written consent of the patient or his/her guardian.

Patients had the option to participate in the study including post-test counseling and the disclosure of test results (consent level A); to participate in the study without post-test counseling (consent level B); or not to participate in the study at all (consent level C). After consent was obtained, the VCT-nurse administered the test. A Determine Rapid HIV1-2 test was carried out on the blood sample and the VCT nurse carried out post-test counseling. Capillus™ HIV-1/HIV-2 confirmatory tests were performed on positive samples, and if the outcomes of both tests were discrepant, a Uni-Gold™ HIV test was done as a tie breaker. Tests were offered free of charge. In total, nine VCT nurses carried out counseling and anywhere between two and four VCT nurses covered each ward. All but one of the counselors was female.

To study whether refusals are more common among patients with a higher likelihood of infection, we use admission diagnosis because it is correlated with HIV status and also observed for patients who were not tested. We use admission rather than discharge diagnosis because it is less likely to be influenced by the test result. All admission diagnoses were coded using the International Classification of Diseases (ICD-10) (35). For each entry we calculated the proportion of HIV positives and these are used as an indicator of the likelihood of infection (table 1). In the analyses that follow, the likelihood of infection is used as a predictor of HIV status for those who refused the test. We therefore assume that HIV prevalence in each group of admission diagnoses is independent from the willingness to be tested. We also use the likelihood of infection variable rather than a set of dummies for the admission diagnosis themselves for purposes of clarity. Substituting one for the other does not change any of the substantive conclusions to be drawn from this study. The pseudo R^2 in a logit regression of HIV status on the likelihood of infection is 0.25, confirming that the latter is a good predictor of HIV status.

Table 1

The likelihood of infection as indicated by the admission diagnosis is first used as a predictor in logistic regressions with the consent level as the outcome of interest. In these models, we verify whether the effect of infection likelihood persists after the inclusion of controls for background characteristics of the respondent as well as features of the study design. In a second step, we estimate bias in HIV prevalence estimates due to refusal via Heckman probit models that account for sample selection. The Heckman sample selection model is a two equation model that consists of a regression equation predicting HIV status ($y = \nu\beta + u_i$), and a selection equation predicting willingness to be tested ($z\gamma + u_2 > 0$) (36-38). The error terms in both equations are assumed to be normally distributed. Ordinary regression estimates are biased when ρ (the correlation between u_1 and u_2) is not zero. The Heckman selection model allows us to use information for

patients who refused the HIV test (i.e. their admission diagnosis and other sociodemographic background characteristics) to improve estimates of parameters in the regression model predicting HIV status.

We limit the study population in four respects. The first set of excluded cases is multiple admissions of the same individuals. We only consider first admissions because higher order admission diagnoses might be influenced by the test outcome at the first visit and thus introduce problems of reverse causality in our models. For the same reason, we exclude individuals who volunteered their HIV status (mostly positive, table 2). The third category of excluded patients is children under 16 years. For infants and children under eighteen months rapid HIV tests cannot distinguish between individual and maternal antibodies. In addition, the dynamics for consent in children may be different than those for adults. The TB/HIV clinic constitutes another special case in our surveillance. HIV testing is standard practice in diagnosing patients in the TB/HIV clinic and some are referred to it precisely for that reason. The TB/HIV clinic of Zewditu hospital was also one of the pioneering facilities for the provision of ART in Ethiopia and this contributes to the (self-)selection of patients into the TB/HIV clinic. We therefore excluded it from the analyses.

4. Results

4.1 Study descriptives

In total 2,719 unique patients were approached for testing. After excluding the TB/HIV clinic and patients under 16 years, 1,650 cases were retained (table 2). Fifty-four of them were discharged prior to being tested and 49 already knew their HIV status. These cases are omitted from further analysis. Of all approached patients 86.10 percent participated in the study (consent A and B), and 75.50 percent chose post-test counseling (consent level A). The percentage of total refusals (13.90 percent, consent level C) is of the same order of magnitude as those observed in the DHS involving serostatus testing in Mali, Kenya and Zambia (5). Of those in consent levels A and B, 29.71 percent tested positive. The share of positives is markedly higher among those not wanting to be informed of their test result (consent level B, 49.55 percent) compared to those that opted for testing and post-test counseling (consent level A, 27.49 percent).

Table 2

4.2 Covariates of consent

Table 3 reports associations between background characteristics and consent. Religion and region of birth are weak predictors of consent and not shown. The age effects are suggestive of an inverse U-shaped pattern with refusals peaking in middle aged adults. Refusal rates are also higher for widows/ers, the divorced and better-educated patients. The most pronounced variability in participation rates was not by

patient characteristics, but by ward and counselor. The first is possibly related to the reason for admission (and hence HIV status), but could be confounded by differential success of counselors in enrolling study participants. Several counselors have refusal rates (consent level C) below 10 percent. For others, the refusal rate varies between 20 and 43 percent. Refusals also declined after the introduction of ART. Particularly relevant for the analysis of bias in HIV-prevalence estimates is the association between the likelihood of infection in the admission diagnosis and refusal: consent for testing and counseling (level A) drops from over 83 percent in patients with the lowest likelihood of infection to just under 70 percent among those with the highest likelihood of infection. To a certain extent, this relationship is counterbalanced by an increasing share of patients who consented to testing without disclosure of results (consent level B) as the likelihood of infection increased.

Table 3

To explore the relationship between refusal and its predictors in a multivariate context, we use logit regression models with the consent level as the outcome of interest (table 4). In the first binary logit model (consent levels B and C versus A), the likelihood of being positive is correlated with refusal for testing and highly significant: for each one percentage point increase in the likelihood of infection, the odds of refusal increase by 1.5 percent. The analysis also confirms that counselors had variable success in obtaining consent. Of further interest is that refusals gradually declined following the introduction of ART. Most of these effects remain stable after the introduction of supplementary controls (model 2). In addition, the odds to consent to testing and counseling are twice as high for women than for men. The effect of age follows the curvilinear pattern described earlier. Those with higher educational status are less likely to participate in testing, which confirms the bivariate results in table 3. In terms of marital status, those who have never been married are most likely to consent.

Table 4

Breaking down the outcome by level of consent (models 3 and 4), changes little in terms of the substantive conclusions compared to the binary logit models. The noteworthy differences are that age is a weak predictor of total refusal (consent level C versus A) and that educational status does not have an effect in the equation predicting consent level B versus A. The parameters for marital status point in the same direction as in the binomial model but vary in their significance level.

4.3 Bias in HIV prevalence estimates

To identify the magnitude of bias introduced by refusal on HIV prevalence estimates, we turn to Heckman probit models of HIV prevalence accounting for sample selection. Heckman regression parameters are used to generate predicted values of HIV prevalence and these are compared with estimates from

standard probit models. All explanatory variables in models 2 and 4 of table 4 are used in the selection equation of the Heckman model. The Heckman regression equation predicting HIV status contains age, a squared term for age, sex, the likelihood of being positive and marital status. These variables are of little substantive interest in this context and are simply chosen to maximize the predictive power of the regression equation. Table 5 presents HIV prevalence estimates based on standard probit models and Heckman probit models under different scenarios. The bottom row reports the likelihood ratio (LR) test for the hypothesis that the error terms in the regression and selection equation are uncorrelated ($H_0: \rho=0$).

In the first column, we present a simple empirical test of the Heckman model. Here we assume we do *not* have information on HIV status of those in consent level B and predict the HIV prevalence in that sample using an ordinary probit model and a probit model that accounts for sample selection (i.e. the Heckman model). Because the HIV prevalence in the total sample (here consent levels A & B) is known, we can compare these estimates with the observed HIV prevalence. The ordinary probit estimate of HIV prevalence is 17.7 percent; the model that accounts for sample selection estimates HIV prevalence at 23.1 percent; and the true or observed value is 22.2 percent. This exercise thus illustrates that Heckman estimates significantly improve upon standard probit estimates of HIV prevalence. The LR test confirms that selection bias is significant. It is noteworthy that ρ is not significantly different from 0 for a Heckman model that only includes basic sociodemographic background characteristics (sex, age, marital status, and education) in the selection equation (not shown). That model also underestimates HIV prevalence. Adding counselor to the selection equation renders ρ significant, but leads to an overestimate of the observed HIV prevalence. Inclusion of information on the health status of patients – an indicator that correlates well with HIV status and consent – thus improves Heckman predictions of HIV prevalence.

Table 5

The last two columns compare HIV estimates for two plausible study designs. The bias in prevalence estimates will be substantial if response is dichotomized into either refusal or full participation without a middle way of testing without disclosure of results (column 2). This scenario is most typical for clinical intervention studies. The bias is much smaller and statistically only marginally significant if the study protocol explicitly allows participants to opt out of post-test counseling. This is shown in the third column. This scenario is more typical for most population-based studies that include HIV serostatus testing.

5. Discussion

While most of the discussion in this paper focused on the relationship between the likelihood of infection and consent for testing, it was not the most important predictor of consent. The largest variation in

consent is accounted for by the counselors, which suggest that studies interested in minimizing non-response must be careful in the selection and training of their fieldwork team. We have no reason to suspect bias in HIV prevalence estimates due to variability in consent attributable to counselors. Another covariate of consent is the availability of ART. In our study, the odds to consent for testing increased by about 20 percent a month following the launch of a governmental ART program in Ethiopia. The absence of a control group, however, does not allow us to exclude other factors that are potentially responsible for this association. The finding that patients are more likely to agree to testing once treatment becomes available is nonetheless plausible and corroborates findings from other studies (39).

The analyses also establish that consent is correlated with the likelihood of HIV infection (measured via the admission diagnosis): patients who participate in the study (consent levels A and B) are less likely to be infected than those who refuse an HIV test (consent level C). This relationship implies that refusal to be tested constitutes a potential source of bias in HIV prevalence estimates. Regression methods that account for sample selection confirm this, but qualification is required in two respects. First, our study is based on a hospital population and still has to be confirmed in a more general sample. A second qualification is that much seems to depend on the study protocol and the informed consent procedures. In this sample, bias is limited if respondents are offered the opportunity to opt-out of post-test counseling. Simply for the sake of scientific accuracy, it is therefore advisable in biomarker collection studies to explicitly provide for that option when introducing the study to the respondent. In studies where the waiting time between testing and feedback is large, this is often a de-facto option. As technological advances reduce the waiting period for test results, however, this will become a consideration of increasing importance. To date, most population-based surveys that included a testing component (e.g., the DHS) have followed a study protocol that did not involve disclosure of HIV status. Our results suggest that in doing so, these surveys have avoided a potentially important source of bias. In health facility-based settings, where the primary concern is often medical intervention rather than epidemiological assessment, the provision to opt out of post-test counseling is not commonly offered, which may help to explain why reported bias in these studies is usually more significant.

While these findings suggest that most population-based surveys follow a study protocol that contains bias, it does not mean that there is no bias at all. First, we also identified a small downward bias in HIV prevalence estimates under the assumptions of a protocol whereby test results are not disclosed. Second, bias may result from a variety of sources other than those studied here. Among these are the limitations of the sampling frame, the refusal rate, other forms of non-response (e.g., population mobility), and probably also the degree to which respondents are aware of their HIV status. If HIV positives are indeed most likely to refuse testing then one should expect that bias in HIV prevalence estimates will be higher in populations

where individuals are more knowledgeable about their HIV status. Because VCT uptake rates may vary considerably across settings and over time, bias in HIV prevalence estimates may vary quite substantially from survey to survey.

6. References

1. **Boerma, J.T. et al.** Measurement of biomarkers in surveys in developing countries: Opportunities and problems. *Population and Development Review*. 27 (2): 303-314 (2001).
2. **WHO/UNAIDS.** Reconciling antenatal clinic-based surveillance and population-based survey estimates of HIV prevalence in Sub-Saharan Africa. Geneva, WHO/UNAIDS, 2003.
3. **Boerma, J.T. et al.** Estimates of HIV-1 prevalence from national population-based surveys as a new gold standard. *Lancet*. 362 (9399): 1929-31 (2003).
4. **Walker, N. et al.** Estimating the global burden of HIV/AIDS: what do we really know about the HIV pandemic? *Lancet*. 363 (9427): 2180-5 (2004).
5. **Garcia-Calleja, J.M. et al.** Lessons learned in the conduct, validation, and interpretation of national population based HIV surveys. *AIDS*. 19 Suppl 2: S9-S17 (2005).
6. **Pison, G. et al.** Seasonal migration: a risk factor for HIV infection in rural Senegal. *J Acquir Immune Defic Syndr*. 6 (2): 196-200 (1993).
7. **Quinn, T.C.** Population migration and the spread of types 1 and 2 human immunodeficiency viruses. *Proc Natl Acad Sci U S A*. 91 (7): 2407-14 (1994).
8. **Nunn, A.J. et al.** Migration and HIV-1 seroprevalence in a rural Ugandan population. *AIDS*. 9 (5): 503-6 (1995).
9. **Decosas, J. et al.** Migration and AIDS. *Lancet*. 346 (8978): 826-8 (1995).
10. **Lurie, M.N. et al.** The impact of migration on HIV-1 transmission in South Africa: a study of migrant and nonmigrant men and their partners. *Sex Transm Dis*. 30 (2): 149-56 (2003).
11. **Lagarde, E. et al.** Mobility and the spread of human immunodeficiency virus into rural areas of West Africa. *Int J Epidemiol*. 32 (5): 744-52 (2003).
12. **Crampin, A.C. et al.** Trends and measurement of HIV prevalence in northern Malawi. *AIDS*. 17 (12): 1817-25 (2003).
13. **Lydie, N. et al.** Mobility, sexual behavior, and HIV infection in an urban population in Cameroon. *J Acquir Immune Defic Syndr*. 35 (1): 67-74 (2004).
14. **Coffee, M.P. et al.** Patterns of movement and risk of HIV infection in rural Zimbabwe. *J Infect Dis*. 191 Suppl 1: S159-67 (2005).
15. **Hull, H.F. et al.** Comparison of HIV-antibody prevalence in patients consenting to and declining HIV-antibody testing in an STD clinic. *JAMA*. 260 (7): 935-8 (1988).
16. **Schwarcz, S.K. et al.** Comparison of voluntary and blinded human immunodeficiency virus type 1 (HIV-1) seroprevalence surveys in a high prevalence sexually transmitted disease clinic population. *Am J Epidemiol*. 137 (6): 600-8 (1993).
17. **Jones, J.L. et al.** HIV seroprevalence and reasons for refusing and accepting HIV testing. *Sex Transm Dis*. 20 (6): 334-7 (1993).
18. **Groseclose, S.L. et al.** Characterization of patients accepting and refusing routine, voluntary HIV antibody testing in public sexually transmitted disease clinics. *Sex Transm Dis*. 21 (1): 31-5 (1994).
19. **Simon, P.A. et al.** Reasons for HIV antibody test refusal in a heterosexual sexually transmitted disease clinic population. *AIDS*. 10 (13): 1549-53 (1996).
20. **Paget, W.J. et al.** Voluntary confidential HIV testing of STD patients in Switzerland, 1990-5: HIV test refusers cause different biases on HIV prevalences in heterosexuals and homo/bisexuals. Swiss Network of Dermatovenereology Policlinics. *Genitourin Med*. 73 (6): 444-7 (1997).
21. **Coulibaly, D. et al.** Attitudes et comportements des femmes enceintes face au dépistage du HIV à Abidjan (Côte D'Ivoire) en 1995 et 1996. Raisons du refus du test et indifférence face aux résultats. *Sante*. 8 (3): 234-8 (1998).
22. **Boxall, E.H. & Smith, N.** Antenatal screening for HIV; are those who refuse testing at higher risk than those who accept testing? *J Public Health (Oxf)*. 26 (3): 285-7 (2004).

23. **Mseleku, M. et al.** HIV seropositive in pregnant South African women who initially refuse routine antenatal HIV screening. *BJOG*. 112 (3): 370-1 (2005).
24. **Meda, N. et al.** HIV infection among pregnant women in Bobo-Dioulasso, Burkina Faso: comparison of voluntary and blinded seroprevalence estimates. *Int J STD AIDS*. 10 (11): 738-40 (1999).
25. **Mpairwe, H. et al.** HIV risk perception and prevalence in a program for prevention of mother-to-child HIV transmission: comparison of women who accept voluntary counseling and testing and those tested anonymously. *J Acquir Immune Defic Syndr*. 39 (3): 354-8 (2005).
26. **Fabiani, M. et al.** Using prevalence data from the programme for the prevention of mother-to-child-transmission for HIV-1 surveillance in North Uganda. *AIDS*. 19 (8): 823-7 (2005).
27. **Obare, F.** The effect of non-response on population-based HIV prevalence estimates: the case of rural Malawi. *Social Networks Project Working Paper*. Philadelphia, University of Pennsylvania, 2005.
28. **Bignami-Van Assche, S. et al.** Evidence from national population-based estimates of bias in HIV prevalence. *Population Association of America Annual Meeting*. Philadelphia 2005.
29. **Mishra, V. et al.** HIV testing in national population-based surveys: experience from the Demographic and Health Surveys. *Bull World Health Organ*. 84 (7): 537-45 (2006).
30. **UNAIDS/WHO.** Guidelines for measuring HIV prevalence in population-based surveys. Geneva, UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance, 2005.
31. **MOH.** AIDS in Ethiopia. Technical Document for the Fifth Edition. Addis Ababa, Ethiopia, Ministry of Health, Disease Prevention and Control Department, 2004.
32. **CSA and ORC Macro.** Ethiopia demographic and health survey 2005. Addis Ababa, Ethiopia and Calverton, Maryland USA, Central Statistical Agency and ORC Macro, 2006.
33. **Araya, T. et al.** Lay diagnosis of causes of death for monitoring AIDS mortality in Addis Ababa, Ethiopia. *Trop Med Int Health*. 9 (1): 178-86 (2004).
34. **ARC.** AIDS Resource Center. HIV Care and ART Update in Ethiopia. Addis Ababa, ARC, 2007.
35. **WHO.** *ICD-10: International statistical classification of diseases and related health problems*. Geneva, World Health Organization, 1993.
36. **Heckman, J.J.** Sample selection bias as a specification error. *Econometrica*. 47 (1): 153-161 (1979).
37. **Winship, C. & Mare, R.D.** Models for sample selection. *Annual Review of Sociology*. 18: 327-350 (1992).
38. **Briggs, D.C.** Causal inference and the Heckman model. *Journal of Educational and Behavioral Statistics*. 29 (4): 397-420 (2004).
39. **Warwick, Z.** The influence of antiretroviral therapy on the uptake of HIV testing in Tutume, Botswana. *International Journal of STD and AIDS*. 17: 470-481 (2006).

Table 1: Admission diagnoses and likelihood of infection. Zewditu Hospital, Addis Ababa (2003-04, age 16+)

	% HIV+	N	ICD-10 code
Diarrhoea and GE of presumed infectious origin	66.67	42	A09
Respiratory TB	69.70	33	A15-16
Other TB	60.00	15	A17-19
HIV	100.00	2	B2
Malaria	17.14	35	B50-54
Herpes zoster, oral candidiasis, toxoplasmosis and PCP	94.74	38	B02, B37, B58-59
Other infectious and parasitic diseases	13.95	43	A01, A03, A07, A30, A35, A41, A63-64, A68, A75, A82, B45
Neoplasm's of breast, cervix, uterus and leiomyoma	14.58	48	C50, C53-55, D25-26
Other neoplasms (benign and malignant)	0.00	25	C0, C2-4, C51-52, C56-58, C6-9, D0, D22-24, D3-4
Thyroid disorders	9.86	71	E00-05
Diabetes and hypoglycemia	11.11	27	E10-E16
Diseases of the nervous system (mainly meningitis)	35.71	14	G00, G03-04, G25, G40, G54
Hypertension	7.14	28	I10-I13
Hypotension	61.90	21	I95
Other diseases of the circulatory system	6.67	45	I05, I09, I15, I21, I31, I38, I49-51, I61, I63-64, I80, I83-I84, I86, I88
Pneumonia	30.56	36	J18
Other diseases of the respiratory system	26.92	26	J11, J44-46, J86, J90, J93-94, J98
Gastritis and other diseases of the oesophagus, stomach and duodenum	15.00	60	K27, K29-31
Diseases of the appendix	8.97	78	K35, K37-38
Hernia and intestinal obstruction	5.71	70	K40, K42-43, K46, K56
Cholelithiasis and diseases of the pancreas	6.06	132	K80, K82, K85-K86
Other diseases of the digestive system	15.79	38	K04, K12, K60, K62-63, K65-66, K72-73, K75-76, K83, K91-93
Diseases of the skin and subcutaneous tissue	22.22	9	L, M
Glomerular diseases and diseases of the urinary system	13.64	22	N0-3
Diseases of male genital organs	2.63	38	N4
Inflammatory diseases of female pelvic organs and disorders of the female genital tract	20.00	25	N7-9
Complications of pregnancy and delivery	15.91	44	O
Fever of unknown origin	32.69	104	R50
Chronic illness	79.31	29	R69
Symptoms signs and abnormal clinical findings not elsewhere specified	17.46	63	R0-4, R56-58, R62
External causes and injuries	7.69	39	S, T, X
Other and unknown admission diagnoses	12.90	31	A80, B19, B56, D5-8, E15, E40-42, E55, E83, E86, E88, K36, P07, Q43, Q53, U, Z4
Total	22.16	1331	

Table 2: Consent for testing and HIV status (Zewditu Hospital, Addis Ababa, 2003-04)

	Freq.	col %	Study participants (col %)	HIV prevalence
Consent level A (testing & post-test counseling)	1,168	70.79	75.50	27.49
Consent level B (testing only)	164	9.94	10.60	49.55
Consent level C (total refusal)	215	13.03	13.90	unknown
Known HIV status	49	2.97	excluded	81.82
Discharged/expired prior to testing	54	3.27	excluded	
Total	1,650	100		

Table 3: Covariates of consent for HIV testing (Zewditu Hospital, Addis Ababa (2003-04))

Age [‡]	Consent level			Total	Counselor [†]	Consent level			Total
	A	B	C			A	B	C	
16-19	85.98	4.67	9.35	107	1	57.53	13.70	28.77	73
20-29	73.90	11.65	14.46	498	2	91.42	4.95	3.63	303
30-39	68.69	14.14	17.17	396	3	53.57	3.57	42.86	28
40-49	77.33	9.72	12.96	247	5	73.17	26.83	0.00	41
50-59	77.46	10.56	11.97	142	6	0.00	0.00	100.0	2
60+	85.90	3.85	10.26	156	7	98.45	0.31	1.24	322
Pearson Chi ² (10) = 28.99	p < .01				8	63.05	16.62	20.33	728
Missing	100.0	0.00	0.00	1	9	56.00	10.00	34.00	50
					Pearson Chi ² (14) = 276.27, p < .01				
Education					Ward				
Illiterate	82.32	10.22	7.46	362	ER	76.16	13.92	9.92	625
1-6 th grade	76.47	11.76	11.76	272	GY	55.22	9.45	35.32	201
7-12 th grade	75.95	11.04	13.01	607	IM	57.73	20.62	21.65	97
>12 th grade	61.24	10.85	27.91	129	SU	84.13	6.09	9.78	624
Pearson Chi ² (8) = 37.46	p < .01				Pearson Chi ² (6) = 134.44, p < .01				
Missing	68.93	7.91	23.16	177					
Marital status					Study month[‡]				
Single	78.59	11.64	9.77	481	Prior to ART	64.04	16.18	19.78	445
Mar	76.46	8.97	14.56	769	Since ART	80.13	8.35	11.52	1,102
Div/wid	69.49	17.8	12.71	118	Pearson Chi ² (2) = 44.72, p < .01				
Pearson Chi ² (4) = 14.49	p < .01								
Missing	67.04	10.06	22.91	179					
Likelihood of infection[‡]					Gender				
≤ 7.49	83.51	6.38	10.11	376	Female	75.34	9.93	14.73	876
7.5 – 14.9	78.97	8.56	12.47	409	Male	75.71	11.48	12.82	671
15.0 – 29.9	69.41	10.64	19.95	376	Pearson chi ² (2)=1.86, p=0.40				
≥ 30	69.87	16.88	13.25	385					
Pearson Chi ² (6) = 44.05	p < .01								
Missing	0.00	100.0	0.00	1					

Notes:

‡ In the regression models that follow, age is defined in terms of single year age groups and study month is coded 0 for the period prior to the introduction of ART and consecutive numbers for months that followed. HIV likelihood is used as the proportion HIV+ for each ICD-10 entry in table 1. The other variables are defined as shown in the table.

† Counselor #4 only worked in the TB/HIV clinic and omitted from this table and any subsequent analysis. Counselor 6 worked primarily in the pediatrics ward.

Table 4: Binary and multinomial logistic regressions predicting refusal of testing for HIV (Zewditu Hospital, Addis Ababa, 2003-04)

	Binary logistic regression predicting refusal (odds ratios)		Multinomial logistic regression predicting refusal (relative risk ratios)			
	B & C versus A		B versus A	C versus A	B versus A	C versus A
	Model 1	Model 2	Model 3	Model 3	Model 4	Model 4
Likelihood of infection	1.01**	1.01**	1.02**	1.01**	1.01**	1.01**
Counselor (vs #1)						
Counselor 2	0.07**	0.07**	0.12**	0.05**	0.06**	0.09**
Counselor 3	0.46	-	0.09**	0.71	-	-
Counselor 5	0.50	0.33	1.58	-	1.48	-
Counselor 6	-	-	-	-	-	-
Counselor 7	0.01**	0.01**	0.01**	0.02**	0.00**	0.01**
Counselor 8	0.46**	0.44	0.56	0.42**	0.26	0.68
Counselor 9	0.65	0.56	0.43	0.78	0.38	0.76
Study month (vs period prior to ART)	0.82**	0.81**	0.78**	0.84**	0.73**	0.88*
Ward (vs ER)						
GY		1.23			0.44	2.68**
IM		1.42			0.56	2.73**
SU		0.83			0.60	1.04
Male		1.96**			1.68**	2.30**
Age		1.05			1.13**	1.02
Age squared		.999*			.998**	1.00
Education (vs no schooling)						
Grade 1-6		1.33			1.16	1.57
Grade 7-12		1.27			0.79	1.95**
> 12 th grade		1.70*			0.82	2.85**
Marital status (vs never married)						
Married		1.44*			1.15	1.71**
Sep/Div//Wid		1.92*			1.78	2.03
N	1544	1357	1546		1359	
LR chi ² (df)	354.57(8)	364.68(18)	406.49(18)		453.04(38)	
Prob > chi ²	0.00	0.00	0.00		0.00	
Pseudo R ²	0.21	0.25	0.18		0.24	
Log likelihood	-680.93	-554.13	-917.05		-732.04	

Notes:

* p ≤ .10; ** p ≤ .05

See table 3 and the notes to that table for a definition of the explanatory variables. Other variables that were controlled for, but omitted in the final models because they lack statistical significance are: birth region (Addis Ababa versus other); religion (Orthodox Christian versus other); a squared term for likelihood of infection; an interaction between the likelihood of infection and study month; an interaction between birth region and sex. Because education and marital status were only introduced as additional variables in the second month of the surveillance, models two and four are based on fewer cases.

Table 5: Comparison of HIV seroprevalence estimates based on standard probit models and models accounting for sample selection under various scenarios (Zewditu Hospital, Addis Ababa, 2003-04)

	Scenario		
	Test of Heckman model [‡]	Post-test counseling is implied	Post-test counseling is optional
E(HIV% - Probit)	17.7 (16.4 -19.1)	17.8 (16.6-19.1)	21.4 (20.2-22.7)
E(HIV% -Heckman)	23.1 (21.7 -24.4)	23.4 (22.1- 24.7)	23.7 (22.4-25.0)
Observed HIV%	22.2 (19.9 – 24.4)	unknown	unknown
Sample	Consent groups A and B	All consent groups	All consent groups
Assumption	HIV status in consent group B is unobserved	HIV status in consent groups B and C is unobserved	HIV status in consent group C is unobserved
LR test $H_0:\rho =0$	$p < .01$	$p < .01$	$p = .07$

Notes: 95%- CI are reported between brackets. Using dummies for admission diagnosis rather than the likelihood of infection in these regressions hardly changes the estimated prevalence rates though one of the selection models did not converge.

[‡] In the first column, we assume that HIV status in consent group B is unknown, and compare the ordinary Probit and Heckman selection model estimate with the true or observed value of HIV prevalence.