AIDS and Behavior (2023) 27:3248–3257
https://doi.org/10.1007/s10461-023-04045-9

ORIGINAL PAPER

# Estimating the Prevalence of over- and Under-Reporting in HIV Testing, Status and Treatment in Rural Northeast South Africa: A Comparison of a Survey and Clinic Records

Hannah H. Leslie[1] · Chodziwadziwa W. Kabudula[2] · Rebecca L. West[3] · Mi-Suk Kang Dufour[4] · Aimée Julien[2,5] · Nkosinathi G. Masilela[2] · Stephen M. Tollman[2] · Audrey Pettifor[2,6] · Kathleen Kahn[2] · Sheri A. Lippman[1,2]

Accepted: 14 March 2023 / Published online: 1 April 2023


## Abstract

We assess the accuracy of self-reported testing, HIV status, and treatment responses compared to clinical records in Ehlanzeni District, South Africa. We linked a 2018 population-based survey of adults 18–49 years old with clinical data at local primary healthcare facilities from 2014 to 2018. We calculated self-reported testing, HIV status, and treatment, and triangulated findings with clinic record data. We adjusted testing estimates for known gaps in HIV test documentation. Of 2089 survey participants, 1657 used a study facility and were eligible for analysis. Half of men and 84% of women reported an HIV test in the past year. One third of reported tests could be confirmed in clinic data within 1 year and an additional 13% within 2 years; these fractions increased to 57% and 22% respectively limiting to participants with a verified clinic file. After accounting for gaps in clinic documentation, we found that prevalence of recent HIV testing was closer to 15% among men and 51% in women. Estimated prevalence of known HIV was 16.2% based on self-report vs. 27.6% with clinic documentation. Relative to clinical records among confirmed clinic users, self report of HIV testing and of current treatment were highly sensitive but non-specific (sensitivity 95.5% and 98.8%, specificity 24.2% and 16.1% respectively), while self report of HIV status was highly specific but not sensitive (sensitivity 53.0%, specificity 99.3%). While clinical records are imperfect, survey-based measures should be interpreted with caution in this rural South African setting.

**Keywords** HIV testing · survey research · self-reported measures · South Africa

## Introduction

The universal test and treat policy for HIV has been incorporated into the South African national HIV/AIDS guidelines since September 2016. The 2017–2022 national strategic plan emphasizes accessible health services (including provider-initiated testing and counseling, community testing, and self-testing) and comprehensive preventive measures [1]. The 2020 national guidelines on HIV management recommend general population testing frequency every 6 to 12 months for sexually active 15–25 year olds and every 12 months for sexually active adults over 25, with more frequent testing for pregnant and breastfeeding women and key populations [2]. Findings from trials of universal testing and treatment strategies suggest that intensive testing is an essential component for such strategies to yield population-level effects on HIV incidence [3]. In an era of highly efficacious antiretroviral therapy (ART), treatment as prevention

✉ Hannah H. Leslie
Hannah.leslie@ucsf.edu

1 Division of Prevention Science, Department of Medicine, University of California, San Francisco, 550 16th Street, 3rd floor, San Francisco, CA 94158, USA

2 MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

3 Boston University School of Public Health, Boston, USA

4 Division of Biostatistics, School of Public Health, University of California, Berkeley, Berkeley, USA

5 Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium

6 Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, USA

(U = U), and pre-exposure prophylaxis (PrEP), broad and frequent testing for HIV is important for both identifying those in need of treatment and linking others to enhanced prevention. To achieve epidemic control in high prevalence settings like South Africa, both high treatment coverage and prevention-effective levels of PrEP will be necessary [4, 5]: routine testing is a precondition for both.

Epidemic monitoring and modeling draw from routine health information systems on HIV testing numbers and positivity rates as well as self-reported testing [6–10]. Aggregate numbers are a poor proxy for population coverage given repeated testing, particularly within reproductive and maternal health services, while self-report of health screenings is often an overestimate [6, 8]. Quantifying the inaccuracy in self-reported measures of HIV service use and status can improve epidemic monitoring and resource allocation. Previous assessments of survey accuracy have focused on self-reported HIV indicators among people living with HIV (PLHIV) compared to biomarker assessment. A 4-country study found sensitivity over 96% for self-reported HIV testing history among PLHIV and > 91% for self-reported awareness of HIV status [11]. A meta-analysis of under-reporting of HIV status found an average of 9% underreporting among general population surveys [12]. However, not all surveys can include biomarkers, refusal rates for biomarker assessment is substantial (39% in the 2017 national survey [13]), and self-reported indicators of HIV testing lack a biomarker for comparison.

There is limited information on self-reported testing history among the general adult population [14], with few if any studies on the extent of mis- or over-reporting HIV testing in particular. Self-reported testing not verified in clinical records could be due to social desirability, misunderstanding, or telescoping (reporting an event as more recent than it was) on the respondent part, data entry error, or incomplete clinical records within and across all available testing sites within the health system. The lack of comprehensive clinical records has limited validation of self-reported testing to date. As HIV testing options continue to expand beyond health facilities, survey-based assessment of testing coverage may increase in importance despite potential inaccuracies. In this study, we evaluate the accuracy of self-reported measures of HIV testing, status, and current treatment comparing responses from a population-based survey with longitudinal clinic records in a rural setting of Ehlanzeni District, South Africa.

## Methods

### Study Setting

This study is based in the Agincourt Health and socio-Demographic Surveillance System (Agincourt HDSS), located in the Bushbuckridge sub-district of Ehlanzeni District, Mpumalanga Province [15]. Ehlanzeni is a national priority district for HIV response [1], with an adult HIV prevalence near 20% [16]. The study site area is part of a former apartheid homeland with high poverty and limited employment opportunities [17]. The Agincourt HDSS was initiated in 1992 and now covers approximately 120,000 residents across 31 villages [15]. At the time of the study, primary healthcare services, including HIV services, were provided by 10 Department of Health (DoH) clinics and community health centers (CHC) – later consolidated to 9 – located within or surrounding the HDSS area (Figure S1 [18]).

### Data Sources

This secondary analysis combines a cross-sectional population-based survey and a longitudinal research database of clinical care records (Figure S2). The survey was conducted from August – December 2018 at the endline of a cluster randomized trial of community mobilization to improve engagement in HIV testing and care [18, 19]. Households with adults 18–49 years old were sampled from the HDSS; individuals resident in the area at least 9 of the past 12 months and consenting to the survey were eligible for study inclusion, regardless of migration status. Of 3305 households successfully contacted, 2182 potentially eligible individuals were identified and 2089 were confirmed eligible and completed the survey (95.7% response rate). The survey included measures of demographics, health behaviors, health care utilization, and HIV knowledge and beliefs.

The Agincourt HDSS-Clinic Link research database was established in 2014 to provide continuous data collection of HIV testing and chronic care visits in local public sector DoH health facilities [18, 20]. Consenting patients provide demographic information used to match patient records to Agincourt HDSS data using a combination of deterministic and probabilistic techniques. Clinical information is extracted from the patient's file at the consenting visit and subsequent visits using a unique ID. However, not all patients consent to HDSS-Clinic Link enrollment (approximately 95% agree) and not all HIV test results are documented in patient files. We conducted a verification exercise in 2017 comparing HDSS-Clinic Link records to clinic testing registers for adults 18–49 and found that HDSS-Clinic Link captured an estimated 73% of tests among women and

58% of tests among men over a 4-month span. Qualitative work within facilities identified that nurses were less likely to start a patient file for individuals testing negative within HIV testing services; files and filing space were prioritized for patients testing positive [21].

We used anonymized IDs to locate survey respondents in HDSS-Clinic Link records from 2013 through December 2018 and performed a manual file check for respondents reporting use of facilities included in the HDSS-Clinic Link but not found by ID.

We limited analysis to users of the 10 DoH facilities included in the HDSS-Clinic Link as of 2017. We defined users of the HDSS-Clinic Link facilities based on survey respondents selecting an HDSS-Clinic Link facility as their source of usual care (if never tested for HIV), most recent HIV test, or HIV care and treatment (if HIV positive), and/or if they had an HIV testing or treatment record active within the past 12 months in the HDSS-Clinic Link system.

### Survey Measures

The survey assessed HIV testing history and testing results in line with global recommendations [22]. Respondents were asked whether they had ever tested for HIV, their most recent test (within the last year, 1–5 years, or more than 5 years ago), test results, and test location. Given prior experience with confusion of the term "HIV test" with clinical measures such as CD4 or viral load count [23], we included clarifying language in the survey, asking, "In your lifetime, how many times have you been tested for HIV? This is the test that tells you if have HIV – it does not mean CD4 or viral load testing. Please also include HIV tests when you did not receive the results." We defined recent testing as reporting a test within the past year. Respondents self-reporting a positive HIV test were asked about diagnosis, care, and ART history. Current treatment was defined as answering yes to questions on ever and still taking ART.

Demographic variables assessed in the survey included age group (18–29, 30–39, 40–49), gender, educational attainment (primary or less, some secondary, secondary, university), marital status, gravidity and current pregnancy status, and earning any income in the past 3 months. The survey included a previously validated scale on anticipated HIV stigma [24, 25]; we dichotomized responses as experiencing any anticipated stigma vs. none.

### HDSS-Clinic Link Measures

We used the HDSS-Clinic Link records to identify date and results of last test prior to date of survey. We defined clinical history of HIV treatment based on having a recorded ART start date and/or having ART prescribed at least once. Clinic-based HIV status was defined based on testing positive and/or any ART use. Current ART use was defined as less than 90 days elapsing between end of the most recent ART prescription and survey date following the DoH definition of loss to follow up. We calculated duration living with HIV as years elapsed from self-reported diagnosis date or first positive test or treatment date in the HDSS-Clinic Link system (whichever came earlier). We excluded individuals aware that they were living with HIV for at least 1 year from analysis of recent HIV testing.

### Administrative data

To triangulate study findings, we extracted the total number of HIV tests from any source and the number of HIV self tests reported in 2018 among adults aged 16 and older by sex from the amfAR PEPFAR Monitoring, Evaluation, and Reporting Database [26]. We extracted 2017 and 2018 population size for women aged 20–49 and men aged 20–49 (the available range closest to study range of 18–49 years) from the 2020 Ehlanzeni District Profile [27].

### Statistical Analysis

The statistical analysis aimed to compare self-report with clinic documentation of HIV testing, HIV status reporting, and HIV treatment, and to assess individual and health system predictors of concordance between self-report and clinic documentation. For HIV testing, we analyzed which of the individuals reporting testing within 12 months had tests confirmed in that time; for HIV status, we assessed self-reporting among PLHIV. We used descriptive statistics to characterize the population eligible for each analysis.

For HIV testing, we cross-tabulated self-reported and clinic-documented testing and calculated percent of self-reported tests that could be confirmed overall and within respondents with an HDSS-Clinic Link file identified. We calculated sensitivity and specificity of self-reported testing limited to individuals with an HDSS-Clinic Link file identified for whom documentation is most likely to be complete. We assessed individual characteristics and testing facility as predictors of confirmed testing within the past year among those reporting a recent test using generalized estimating equation (GEE) models clustered by village with a logit link and exchangeable correlation. We repeated these models limited to individuals with an HDSS-Clinic Link file as a sensitivity analysis.

We calculated population prevalence of recent testing based on self-report and using only tests confirmed in the HDSS-Clinic Link system, overall and separately for men and women. We corrected estimates of confirmed tests to account for under-documentation from the 2017 verification exercise using the following steps:

1. Creating an indicator equal to 1 for all individuals with confirmed testing plus a randomly selected subset of individuals with unconfirmed test to completely recapture the tests truly conducted but not documented based on the verification exercise (27% of tests among women, 42% among men); the over-imputed indicator approximates true testing in a counterfactual setting of complete documentation.
2. Modeling the over-imputed indicator in GEE models separately for men and women as well as overall, adjusting for facility.
3. Predicting prevalence of confirmed testing based on these models of the over-imputed indicator with facility set to the reference facility - a CHC with large patient population – to standardize across facilities.
4. Repeating steps 1–3 1000 times.
5. Combining results to calculate prevalence of confirmed testing with 95% confidence intervals using Rubin's rule [28].

As an external validity check of testing frequency, we calculated per capita testing by dividing total tests reported for all adults (grouped as 16 and older) in the PEPFAR monitoring system by population aged 20–49 in Ehlanzeni District, overall and stratified for men and women. This estimate is imperfect as it is an aggregate figure and the age ranges do not align exactly. It can be interpreted as an upper bound on testing frequency among adults 18–49 given that the numerator is the number of tests, including repeated tests among individuals such as pregnant and breastfeeding women tested multiple times, and includes tests in adults over 50, neither of which are captured in the denominator.

For HIV status, we calculated prevalence of known HIV in the study population based first on self-report alone and then on self-report or clinic documentation of positive test or ART history. We cross tabulated self-report and clinic documented status and calculated sensitivity and specificity of self-reported status among those with HDSS-Clinic Link files found, classifying those with unknown or unstated status as negative. Among all PLHIV, we assessed individual characteristics and anticipated stigma associated with self-report of living with HIV using GEE models as described above.

Finally, to assess current ART use, we calculated prevalence of treatment among those self-reporting HIV positive status. We cross-tabulated self-report with clinic documentation and calculated sensitivity and specificity among those with HDSS-Clinic Link files found.

Descriptive analyses incorporate survey sampling weights; cross-tabulations present unweighted counts and GEE analyses are unweighted.

## Results

Of the 2089 respondents with completed surveys, 1657 individuals were eligible for this analysis based on use of an HDSS-Clinic Link facility. The 432 excluded individuals were more likely to be male, have secondary or higher educational attainment, be unmarried, self-report not living with HIV, and to identify private health services or non-Clinic Link DoH facilities (such as the nearest hospital) as their usual source of care (Table S1).

The analytic sample included more women than men (62.6% female, Table 1). Most had received at least some (46.1%) or completed (36.4%) secondary education, were unmarried (59.2%), and had not earned income in the past 3 months (56.7%). Nearly all reported a study facility as their usual source of care, and HDSS-Clinic Link records were identified for over half of respondents. Anticipated stigma was low, with only one third of respondents anticipating any stigma if living with HIV; 65 women (6.6%) reported being pregnant at the time of the survey.

### HIV Testing

Of the 1272 individuals eligible for HIV testing, over two thirds of respondents reported a recent HIV test (84% of women and 50% of men); other characteristics of those reporting testing were similar to all eligible for testing.

Of the 880 individuals reporting a recent HIV test, 98% reported testing within an HDSS-Clinic Link facility; 294 (33.4%, 95% confidence interval [CI] 30.4%, 36.6%) were matched to a test in HDSS-Clinic Link records in the same time frame and an additional 13% had an older test on record (Table 2 A). The remaining 54% of self-reported recent tests could not be confirmed in clinic files. Excluding individuals who were not found in HDSS-Clinic Link files at all, 57% of recent tests could be confirmed within 1 year (95% CI 52.6%, 61.1%) and an additional 22% (95% CI 18.9%, 26.0%) more than 1 year prior. We found little evidence of under-reporting of testing: only 3 individuals reporting never testing had a confirmed test in clinical documentation. Restricting to those with clinical files found, self-reported testing within the past 12 months had a sensitivity of 95.5% (95% CI 92.5%, 97.3%) and specificity of 24.2% (95% CI 19.6%, 29.4%).

Confirmed tests were more common among women, particularly currently pregnant women, than men, and in individuals earning any income. Odds of confirmed test differed

**Table 1** Demographics of analytic sample by eligibility for each analysis, survey weighted

| | | HIV testing | | HIV status and ART | | |
|---|---|---|---|---|---|---|
| | All respondents | Eligible for HIV testing | Self-report recent test | Living with HIV, self-report and/or documentation | Self-report living with HIV | Self-report current ART |
| N | 1657 | 1272 | 880 | 424 | 250 | 238 |
| Weighted % of total | 100 | 74.5 | 51.7 | 27.7 | 16.2 | 15.7 |
| Gender | | | | | | |
| Male | 621 (37.4%) | 557 (43.7%) | 276 (31.3%) | 85 (19.8%) | 44 (17.6%) | 42 (17.3%) |
| Female | 1037 (62.6%) | 716 (56.3%) | 605 (68.7%) | 340 (80.2%) | 207 (82.4%) | 197 (82.7%) |
| Age categories | | | | | | |
| 18–29 | 726 (43.8%) | 666 (52.4%) | 445 (50.5%) | 88 (20.6%) | 44 (17.3%) | 38 (15.9%) |
| 30–39 | 472 (28.5%) | 322 (25.2%) | 249 (28.3%) | 159 (37.5%) | 99 (39.5%) | 97 (40.6%) |
| 40–49 | 460 (27.7%) | 286 (22.4%) | 187 (21.2%) | 178 (41.9%) | 109 (43.3%) | 104 (43.5%) |
| Education | | | | | | |
| Primary or less | 207 (12.4%) | 145 (11.4%) | 84 (9.5%) | 67 (15.7%) | 42 (16.5%) | 39 (16.4%) |
| Some secondary | 764 (46.1%) | 573 (45.0%) | 371 (42.1%) | 208 (48.9%) | 129 (51.4%) | 123 (51.3%) |
| Secondary | 604 (36.4%) | 485 (38.1%) | 370 (42.0%) | 134 (31.5%) | 74 (29.5%) | 71 (29.7%) |
| University | 85 (5.1%) | 70 (5.5%) | 57 (6.4%) | 17 (3.9%) | 7 (2.6%) | 7 (2.7%) |
| Marital status | | | | | | |
| Never married | 981 (59.2%) | 804 (63.2%) | 531 (60.3%) | 206 (48.6%) | 123 (49.0%) | 116 (48.7%) |
| Married | 522 (31.5%) | 394 (30.9%) | 307 (34.9%) | 140 (32.9%) | 92 (36.4%) | 87 (36.2%) |
| Separated or widowed | 155 (9.3%) | 75 (5.9%) | 43 (4.8%) | 79 (18.5%) | 37 (14.6%) | 36 (15.1%) |
| Currently pregnant | | | | | | |
| No | 908 (93.4%) | 625 (92.6%) | 529 (91.6%) | 307 (93.8%) | 180 (91.2%) | 173 (91.2%) |
| Yes | 65 (6.6%) | 50 (7.4%) | 49 (8.4%) | 21 (6.2%) | 18 (8.8%) | 17 (8.8%) |
| Income past 3 months | | | | | | |
| No | 941 (56.7%) | 761 (59.8%) | 514 (58.4%) | 200 (47.1%) | 116 (46.0%) | 109 (45.7%) |
| Yes | 717 (43.3%) | 512 (40.2%) | 367 (41.6%) | 225 (52.9%) | 135 (54.0%) | 130 (54.3%) |
| Usual source of care | | | | | | |
| Clinic Link facility | 1614 (97.4%) | 1231 (96.8%) | 848 (96.3%) | 421 (99.2%) | 249 (99.5%) | 237 (99.5%) |
| Other DoH facility | 21 (1.3%) | 19 (1.5%) | 14 (1.5%) | 3 (0.6%) | 1 (0.2%) | 1 (0.3%) |
| Private/other | 23 (1.4%) | 23 (1.8%) | 20 (2.2%) | 2 (0.2%) | 1 (0.2%) | 1 (0.2%) |
| Anticipated stigma (any) | | | | | | |
| None | 1102 (66.5%) | 851 (66.9%) | 601 (68.3%) | 275 (64.7%) | 163 (65.2%) | 154 (64.5%) |
| Any | 556 (33.5%) | 422 (33.1%) | 280 (31.7%) | 150 (35.3%) | 88 (34.8%) | 85 (35.5%) |
| Clinic Link file | | | | | | |
| Not found | 666 (40.1%) | 648 (50.9%) | 350 (39.7%) | 44 (10.3%) | 45 (17.6%) | 40 (16.6%) |
| Found | 992 (59.9%) | 625 (49.1%) | 531 (60.3%) | 381 (89.7%) | 206 (82.4%) | 199 (83.4%) |
| Years living with HIV (Median, IQR) | NA | NA | NA | 4.5 (2.0, 7.3) | 5.0 (3.0, 8.3) | 5.5 (3.0, 8.3) |

*Numbers may sum to > N due to rounding of survey-weighted count.

across testing facilities (Table 3). In adjusted models, adults 30–39 and 40–49 also had lower odds of confirmed tests as compared to younger adults. Individuals reporting any anticipated stigma had higher odds of confirmed test in adjusted analysis. Results were similar in sensitivity analysis among only those with Clinic Link files except adjusted odds of confirmed test did not differ significantly by age (Table S2).

Overall, self-reported recent testing was 69% (95% CI 66%, 72%); confirmed testing for the same time frame was estimated as only 23% (95% CI 20.9%, 25.5%), increasing to 34% (95% CI 29%, 39%) when accounting for under-documentation in clinical files (Fig. 1). Self-reported testing was twice as high as confirmed testing accounting for under-documentation overall, 1.67 times higher for women (86% vs. 51%), and over 3 times higher for men (50% vs. 15%).

PEPFAR reporting indicated that 432,955 HIV tests were conducted among adults in Ehlanzeni District in 2018 (274,923 for women, 158,032 for men); no self-tests were reported in this calendar year. Given a population of 1,832,551 adults aged 20–49, the upper bound on population-level testing would be 56% (67% for women, 44% for

**Table 2** Comparison of self-report to clinic documentation for HIV testing, status, and current ART

| **A: HIV testing among those not previously positive (N = 1272)** | | | | | | |
|---|---|---|---|---|---|---|
| | | **Testing in HDSS-Clinic Link records** | | | | |
| | | Participant file found | | | No file found | **Total** |
| | | ≤ 12 months | > 12 months | No record of test | | |
| **Self-reported testing** | ≤ 12 months | 294 | 115 | 108 | 363 | 880 |
| | > 12 months | 12 | 23 | 40 | 205 | 280 |
| | Never | 2 | 1 | 7 | 102 | 112 |
| | **Total** | 308 | 139 | 155 | 670 | 1272 |
| **B: HIV status among all respondents (N = 1657)** | | | | | | |
| | | **Status in HDSS-Clinic Link records** | | | | |
| | | Participant file found | | | No file found | |
| **Self-reported status** | | HIV+ | HIV- | No record of status | | **Total** |
| | HIV+ | 196 | 3 | 2 | 49 | 250 |
| | HIV- | 146 | 403 | 146 | 545 | 1240 |
| | Unknown / declined[1] | 28 | 12 | 12 | 115 | 167 |
| | **Total** | 370 | 418 | 160 | 709 | 1657 |
| **C: ART among those reporting living with HIV (N = 250)** | | | | | | |
| | | **ART use in HDSS-Clinic Link records** | | | | |
| | | Participant file found | | | No file found | |
| | | Current | Lapsed | No record of ART | | **Total** |
| **Self-reported ART use** | Current | 168 | 23 | 3 | 44 | 238 |
| | Lapsed | 0 | 2 | 1 | 0 | 3 |
| | Never | 2 | 1 | 1 | 5 | 9 |
| | **Total** | 170 | 26 | 5 | 49 | 250 |

1 Includes never tested, tested and did not receive results, and declined to state results.

men), assuming no individuals tested more than once and no individuals over 50 tested.

## HIV Status

Of 1657 respondents, 424 were classified as persons living with HIV (PLHIV) based on self-report and/or clinic documentation (Table 1, column 5). As expected based on the epidemic in this setting, PLHIV were older and more likely to be women than those not living with HIV. Median duration since diagnosis was 4 years. The 250 individuals (59%) self-reporting living with HIV shared similar demographics to all PLHIV, though median duration living with HIV was longer (5 years) among these respondents. Incorporating survey weighting, 16.2% (95% CI 14.2%, 18.4%) of study respondents were known to be living with HIV based on self-report alone and 27.6% (95% CI 25.1%, 30.2%) including clinic documentation. Prevalence of known HIV among women would be 21.4% on self-report vs. 35.5% including unreported diagnoses, and for men 7.5% on self-report alone nearly doubling to 14.5% with unreported diagnoses confirmed in clinic files.

The majority of individuals self-reporting living with HIV had HDSS-Clinic Link files with records of positive test and/or ART history: 196 of 250 (78.4%, 95% CI 72.9%, 83.0%, Table 2B). Five individuals with HDSS-Clinic Link files had either negative (N = 3) or no recorded HIV results, and the remaining 49 individuals were not found in HDSS-Clinic Link files. Based on clinic records, 11.7% (95% CI 10.1%, 13.7%) of those reporting negative test results and 16.7% (95% CI 11.8%, 23.2%) of those not reporting results on the survey were living with HIV. Nearly all (403 of 418) individuals confirmed HIV negative in clinic documentation reported their status as negative, while 12 reported unknown status or declined to state results. Limiting to those with HDSS-Clinic files and HIV status on record, sensitivity of self-reported HIV status was 53.0% (95% CI 47.9%, 58.0%) and specificity was 99.3% (97.8%, 99.8%).

In unadjusted analysis, adults 40–49 and those with longer time since HIV diagnosis had higher odds of self-reporting their HIV positive status, while those with university education were less likely to report compared to individuals with primary or less education (Table 4). In the adjusted model, only years since HIV diagnosis was statistically significantly associated with self-report of HIV status among PLHIV (AOR 1.11, 95% CI 1.05, 1.18). Results were unchanged in sensitivity analysis limited to those with HDSS-Clinic Link files (Table S3).

## ART History

Nearly all (238 of 250, 95.2%) of those self-reporting HIV diagnosis indicated current ART use (Table 2 C); we identified any ART record for 191 (80.3%) of them, of which

**Table 3** Association of individual characteristics with confirmed test among those reporting recent HIV test (N = 880)

| | Unadjusted | Adjusted |
|---|---|---|
| | OR (95% CI) | AOR (95% CI) |
| Gender | | |
| Male | REF | REF |
| Female, not pregnant | **3.89 (2.71, 5.59)** | **4.41 (3.01, 6.44)** |
| Female, pregnant | **11.07 (5.70, 21.51)** | **13.26 (6.63, 26.52)** |
| Age | | |
| 18–29 | REF | REF |
| 30–39 | 0.90 (0.65, 1.25) | **0.68 (0.47, 0.99)** |
| 40–49 | 0.75 (0.50, 1.12) | **0.58 (0.36, 0.94)** |
| Education | | |
| Primary or less | REF | REF |
| Some secondary | 0.86 (0.48, 1.53) | 1.00 (0.51, 1.95) |
| Secondary | 0.86 (0.48, 1.53) | 0.81 (0.42, 1.59) |
| University | 1.17 (0.56, 2.45) | 1.33 (0.58, 3.07) |
| Recent income (any) | **1.44 (1.09, 1.92)** | **1.57 (1.14, 2.18)** |
| Anticipated stigma (any) | 1.34 (0.99, 1.80) | **1.45 (1.05, 2.02)** |
| Health facility | | |
| CHC 1 | REF | REF |
| CHC 2 | 0.75 (0.45, 1.25) | 0.61 (0.35, 1.07) |
| Clinic 1 | 1.15 (0.72,1.84) | 1.09 (0.63, 1.87) |
| Clinic 2 | 0.92 (0.55,1.54) | 1.02 (0.57, 1.83) |
| Clinic 3 | **0.46 (0.24, 0.85)** | **0.44 (0.22, 0.87)** |
| Clinic 4 | 0.62 (0.38, 1.04) | 0.68 (0.38, 1.20) |
| Clinic 5 | **0.49 (0.33, 0.73)** | **0.39 (0.25, 0.60)** |
| Clinic 6 | 1.22 (0.67, 2.22) | 1.20 (0.62, 2.34) |
| Clinic 7 | **0.29 (0.17, 0.50)** | **0.21 (0.12, 0.39)** |
| Clinic 8 | 1.30 (0.70, 2.40) | 1.31 (0.65, 2.62) |

**Table 4** Association of individual characteristics with disclosure of HIV positive status (N = 424)

| | Unadjusted | Adjusted |
|---|---|---|
| | OR (95% CI) | AOR (95% CI) |
| Gender | | |
| Male | REF | REF |
| Female | 1.33 (0.86, 2.08) | 1.45 (0.91, 2.31) |
| Age | | |
| 18–29 | REF | REF |
| 30–39 | 1.49 (0.91, 2.44) | 1.17 (0.69, 1.98) |
| 40–49 | **1.75 (1.05, 2.90)** | 1.22 (0.68, 2.18) |
| Education | | |
| Primary or less | REF | REF |
| Some secondary | 1.10 (0.62, 1.95) | 1.19 (0.65, 2.19) |
| Secondary | 0.70 (0.38, 1.29) | 0.80 (0.42, 1.51) |
| University | **0.30 (0.09, 0.93)** | 0.34 (0.10, 1.13) |
| Recent income (any) | 1.14 (0.78, 1.67) | 1.12 (0.75, 1.68) |
| Anticipated stigma (any) | 1.07 (0.70, 1.62) | 1.17 (0.76, 1.81) |
| Years since HIV diagnosis | **1.13 (1.06, 1.19)** | **1.11 (1.05, 1.18)** |

168 (88.0%) indicated active treatment. Incorporating clinic documentation, current ART use among those self-reporting HIV diagnosis would be 68% overall or 87.6% limited to those with an identified HDSS-Clinic Link file. Three of nine respondents indicating no lifetime use of ART had a treatment record, including two actively on ART. Among those with HDSS-Clinic Link files, sensitivity of reporting current ART use was 98.8% (95% CI 95.4%, 99.7%) and specificity was 16.1% (95% CI 6.9%, 33.4%).

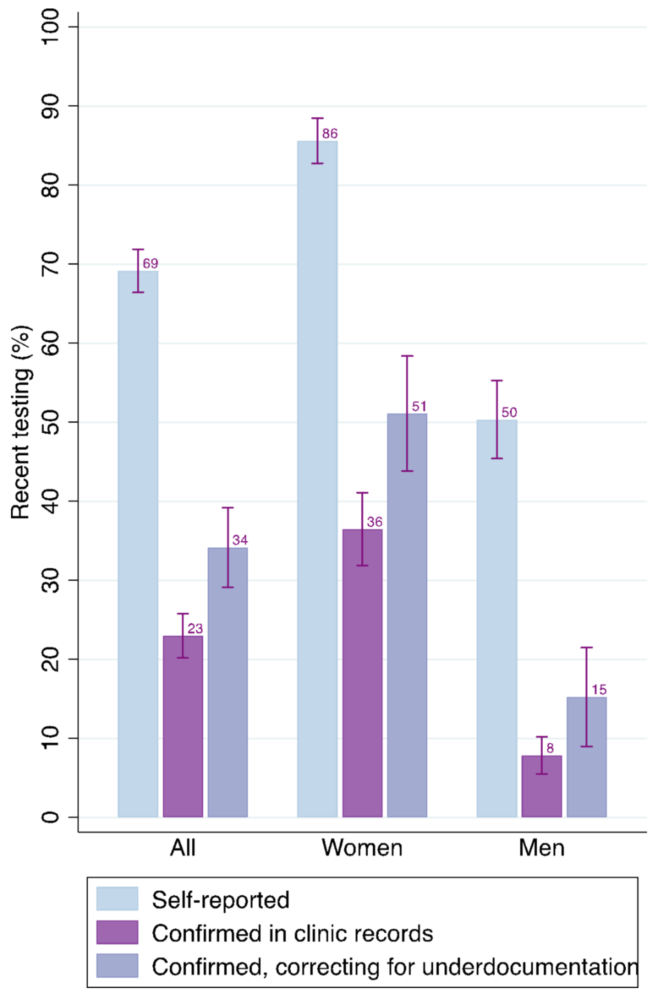


**Fig. 1** Prevalence of HIV testing within 12 months based on self-report and with clinic record confirmation

## Discussion

By combining a population-based survey with health system documentation, we found self-reported HIV testing in the past 12 months to be twice as high as confirmed clinic testing. Among men, self-reported testing was 3 times greater than confirmed testing, even after accounting for underdocumentation in clinic records. Accounting for underdocumentation, we estimated 51% of women and 15% of men definitely tested for HIV within the past year, likely insufficient coverage for effective prevention and treatment efforts in this high-incidence area, and far lower than the survey-based estimates of 84% of women and 50% of men, which are comparable to self-reported testing nationally [29]. Among individuals known to be living with HIV in this study, only half self-reported positive status, resulting in a 10%-point underestimate of prevalence of known HIV. Self report overestimated current ART use by at least 7% points. Type of measure is critical in assessing self-reported measures: self-report of testing and treatment were highly sensitive but non-specific (false negatives were rare but false positives common), while self report of HIV status was highly specific but not sensitive. As a whole, these findings suggest caution in using survey-based, self-reported measures of HIV testing and HIV status in epidemic monitoring, including for the 95-95-95 targets. Most notably, true testing levels may be somewhat higher than clinic-based estimates and substantially lower than survey-based estimates, especially for men.

Our estimates of confirmed testing within 12 months are higher than purely clinic-based estimates for the whole population [19] – reflecting the selection of non-migrants and users of HDSS-Clinic Link facilities for this analysis as well as the corrections we applied for documentation gaps – and much lower than self-reported figures. Comparing the estimate of 69% adults testing on self report to the liberal upper bound of 56% calculated from total adult tests over the 20–49 year old population in Ehlanzeni District in 2018 underscores the discrepany in self-reported testing prevalence.

Multiple factors may contribute to discrepancies between self-reported testing history and clinic-confirmed testing. Respondent error in reporting a recent test could be due to social desirability, recall error on the time or place of testing, misinterpretation of the question, or data entry error, as found in studies of HIV status reporting [23]. Prior studies have investigated stigma as a cause of misreporting, particularly underreporting [11]. We found few instances of individuals underreporting a documented HIV test and no evidence that stigma was linked to more erroneous reporting in this setting. Some respondents could have accessed testing outside of HDSS-Clinic Link facilities through private clinics or Agincourt HDSS research studies [30, 31], although only 2% reported using such sources; private care and self-testing were scarce and expensive at the time of the study. Gaps in health system documentation are a second important factor for discrepancies, as confirmed during our verification exercise and in our finding of notable between-clinic differences in confirmed testing; prior research suggests negative tests are less likely to be documented [32]. In our study, pregnant women were most likely to have tests verified, potentially reflecting their greater engagement with the health system. Third, some survey respondents truly receiving services at HDSS-Clinic Link facilities may not have been found in HDSS-Clinic Link files due to not consenting to data capture or not having a linked census ID; this is the most likely explanation for the 18% of respondents reporting ART use who could not be linked to files. While under-documentation within the health system is substantial and surely accounts for some of the observed discrepancy, the apparent overreporting of recent testing remained 2-fold - and 3-fold for men – even after analysis to correct for incomplete records.

In contrast to reported testing, reported HIV status based on survey estimates was a 10-percentage point underestimate of prevalence of known HIV based on both self-report and clinic documentation. Sensitivity of self-report was only 53% compared to clinic documentation, while specificity exceeded 99%; the small number of apparent false positive reports of HIV status may suggest respondent misunderstanding or data entry error. While underreporting of HIV status was more substantial in this study than a prior meta-analysis,[12] our findings are quite similar to a recent study of nearly 2000 adults in the Agincourt DHSS area completing dried blood spot HIV testing from 2018–2019.[31] In that study, Yorlets et al. found HIV prevalence of 25.3% in their analytic sample along with 43.9% sensitivity and 99.0% specificity of self-reported status. Sensitivity may have appeared higher in our study as we did not have blood spot testing to identify undiagnosed infections. In keeping with findings the recent study in the same setting,[31] we found that individuals living with HIV for longer were more likely to report it, perhaps suggesting greater comfort with disclosing HIV status over time. Survey report of HIV status among PLHIV did not differ by gender or age in our study, although prior studies have found lower accuracy in self-reported status among younger respondents and at times among men [11, 31, 33, 34]. Our findings suggest substantial misreporting despite the use of items based on global recommendations for questions for HIV status and history and do not identify major demographic predictors of inaccuracy. These findings call for greater consideration of validation efforts such as repeated surveys of a subset of respondents or alternative formulations of questions on HIV testing and status to allow for analytic correction of resulting estimates.

Self-reported current treatment (95%) exceeded confirmed active treatment (88% among those with records found), although the difference was not large. Social desirability or differences in interpreting "still taking" ART could account for the 10% of respondents reporting currently taking ART whose prescription had lapsed more than 90 days prior.

This study drew from the well defined HDSS census area, a setting where the DoH facilities included in the HDSS-Clinic Link research database are the main source of health services, enabling us to compare self-report to clinical care for testing as well as HIV and treatment status. We supplemented clinical records with information on gaps in documentation to adjust estimates of confirmed test records and compared to aggregate data to bound our estimates; we repeated explanatory models limiting to those with files identified as a robustness check. The study findings are limited by the lack of gold standard comparison to calculate accuracy: clinic records are incomplete particularly for HIV testing, and the survey did not include biomarker assessment of HIV and ART status. Estimates of HIV prevalence and the sensitivity and specificity of self-report reflect known HIV status, not population-based biomarker testing. The population included for analysis is not fully generalizable due to differences between this population and non-users of DoH facilities, who were more likely to be male, unmarried, and more highly educated.

## Conclusion

Surveys provide an essential complement to clinical records in providing population-based estimates of testing, status, and treatment uptake, particularly for individuals who access differentiated service models, change sources of care, or seek care outside the government health system. Improving survey design to address overreporting of testing and current ART as well as underreporting of HIV status in this setting is critical for accurate monitoring of population testing and of progress towards the 95-95-95 goals. Our findings that 51% of women and 15% of men definitely testing in the past year underscore the challenges in reaching adequate levels of population HIV testing even before the Covid-19 pandemic and the renewed attention required to ensure broader testing efforts, particularly for men.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10461-023-04045-9.

**Authors' Contributions** Conceptualization: Hannah H Leslie, Chodziwadziwa W. Kabudula, Sheri A Lippman. Methodology: Hannah H Leslie, Chodziwadziwa W. Kabudula, Sheri A Lippman. Formal analysis: Hannah H Leslie, Chodziwadziwa W. Kabudula. Investigation: Chodziwadziwa W. Kabudula, Rebecca L West, Aimée Julien, Nkosinathi G Masilela, Stephen M Tollman, Audrey Pettifor, Kathleen Kahn, Sheri A Lippman. Writing – original draft: Hannah H Leslie. Writing – review and editing: All authors. Visualization: Hannah H Leslie. Supervision: Sheri A Lippman. Project administration: Sheri A Lippman. Funding acquisition: Stephen M Tollman, Audrey Pettifor, Kathleen Kahn, Sheri A Lippman.

**Funding** This research is supported by the United States National Institute of Mental Health (R01MH103198; Lippman/Pettifor) and by a grant from the National Institutes of Health, UCSF-Bay Area Center for AIDS Research, P30AI027763. Supplementary funding for grant activities was received in the second year by PEPFAR through Right to Care (USAID Cooperative Agreement AID 674-A-12-000020) and Project SOAR (USAID Cooperative Agreement AID-OAA-14-00060) mechanisms. The Agincourt HDSS is supported by the South African Medical Research Council and University of the Witwatersrand, the Dept. of Science and Innovation, as well as (previously) the Wellcome Trust, UK (grants 058893/Z/99/A; 069683/Z/02/Z; 085477/Z/08/Z; 085477/B/08/Z). The funding institutions have not participated in the design or implementation of the study.

**Data Availability** Data used for analysis are available from the authors on reasonable request.

**Code Availability** Code written for this analysis is available from the authors on reasonable request.

## Declarations

**Conflicts of Interest** Dr. Leslie reports research funding from the National Institutes of Health, the Bill & Melinda Gates Foundation, and the InterAmerican Development Bank during the course of this research. The authors report no conflicts of interest.

**Ethics approval** All data collection procedures were approved by the Human Research Ethics Committee (Medical) at the University of the Witwatersrand, Mpumalanga Department of Health and Social Development Research and Ethics Committee, and Institutional Review Boards at the University of North Carolina, Chapel Hill and the University of California, San Francisco.

**Consent to participate** Participants provided written informed consent at enrollment into studies providing data for this secondary analysis.

**Consent for publication** Not applicable.

## References

1. Let Our Actions Count. : South Africa's National Strategic Plan for HIV, TB and STIs 2017–2022 [Internet]. Pretoria, South Africa: National Department of Health; 2017 May [cited 2022 Apr 4]. Available from: https://www.gov.za/sites/default/files/gcis_document/201705/nsp-hiv-tb-stia.pdf
2. National Department of Health. National Consolidated Guidelines for the Management of HIV in Adults, Adolescents, Children and Infants and Prevention of Mother-to-Child Transmission [Internet]. South Africa: National Department of Health. ; 2020 Feb [cited 2023 Jan 31]. Available from: https://sahivsoc.org/Files/National%20Consolidated%20Guidelines%2030062020%20signed%20PRINT%20v7.pdf
3. Hayes RJ, Donnell D, Floyd S, Mandla N, Bwalya J, Sabapathy K et al. Effect of Universal Testing and Treatment on HIV Incidence — HPTN 071 (PopART).New England Journal of Medicine. 2019 Jul18;381(3):207–18.
4. Akullian A, Morrison M, Garnett GP, Mnisi Z, Lukhele N, Bridenbecker D et al. The effect of 90-90-90 on HIV-1 incidence and mortality in eSwatini: a mathematical modelling study. The Lancet HIV. 2020 May 1;7(5):e348–58.
5. Baggaley RF, Hollingsworth TD. How universal does universal test and treat have to be? The Lancet HIV. 2020 May 1;7(5):e306–8.
6. Giguère K, Eaton JW, Marsh K, Johnson LF, Johnson CC, Ehui E et al. Trends in knowledge of HIV status and efficiency of HIV

testing services in sub-Saharan Africa, 2000–20: a modelling study using survey and HIV testing programme data. The Lancet HIV. 2021 May 1;8(5):e284–93.
7. Johnson LF, van Rensburg C, Govathson C, Meyer-Rath G. Optimal HIV testing strategies for South Africa: a model-based evaluation of population-level impact and cost-effectiveness. Sci Rep. 2019 Sep 2;9(1):12621.
8. Johnson LF, Rehle TM, Jooste S, Bekker LG. Rates of HIV testing and diagnosis in South Africa: successes and challenges. AIDS. 2015 Jul;17(11):1401–9.
9. Johnson LF, Dorrington RE, Moolla H. Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa. South Afr J HIV Med. 2017 Jul;27(1):694.
10. Maheu-Giroux M, Marsh K, Doyle CM, Godin A, Lanièce Delaunay C, Johnson LF et al. National HIV testing and diagnosis coverage in sub-Saharan Africa: a new modeling tool for estimating the 'first 90' from program and survey data.AIDS. 2019 Dec15;33(Suppl 3):S255–69.
11. Xia Y, Milwid RM, Godin A, Boily MC, Johnson LF, Marsh K et al. Accuracy of self-reported HIV-testing history and awareness of HIV-positive status in four sub-Saharan African countries. AIDS. 2021 Mar 1;35(3):503–10.
12. Soni N, Giguère K, Boily MC, Fogel JM, Maheu-Giroux M, Dimitrov D et al. Under-Reporting of Known HIV-Positive Status Among People Living with HIV: A Systematic Review and Meta-analysis. AIDS Behav. 2021 Dec 1;25(12):3858–70.
13. Marinda E, Simbayi L, Zuma K, Zungu N, Moyo S, Kondlo L, et al. Towards achieving the 90-90-90 HIV targets: results from the south African 2017 national HIV survey. BMC Public Health. 2020 Sep;9(1):1375.
14. Rentsch CT, Reniers G, Machemba R, Slaymaker E, Marston M, Wringe A et al. Non-disclosure of HIV testing history in population-based surveys: implications for estimating a UNAIDS 90-90-90 target.Glob Health Action. 2018 Dec14;11(1):1553470.
15. Kahn K, Collinson MA, Gómez-Olivé FX, Mokoena O, Twine R, Mee P et al. Profile: Agincourt Health and Socio-demographic Surveillance System. Int J Epidemiol. 2012 Aug 1;41(4):988–1001.
16. Gómez-Olivé FX, Angotti N, Houle B, Klipstein-Grobusch K, Kabudula C, Menken J, et al. Prevalence of HIV among those 15 and older in rural South Africa. AIDS Care. 2013;25(9):1122–8.
17. Twine R, Kahn K, Scholtz A, Norris SA. Involvement of stakeholders in determining health priorities of adolescents in rural South Africa. Global Health Action. 2016 Dec 1;9(1):29162.
18. Lippman SA, Pettifor A, Rebombo D, Julien A, Wagner RG, Kang Dufour MS, et al. Evaluation of the Tsima community mobilization intervention to improve engagement in HIV testing and care in South Africa: study protocol for a cluster randomized trial. Implement Sci. 2017 Jan;17(1):9.
19. Lippman SA, Pettifor A, Dufour MSK, Kabudula CW, Twine R, Peacock D et al. A community mobilisation intervention to improve engagement in HIV testing, linkage to care, and retention in care in South Africa: a cluster-randomised controlled trial. The Lancet HIV. 2022 Sep 1;9(9):e617–26.
20. Kabudula CW, Clark BD, Gómez-Olivé FX, Tollman S, Menken J, Reniers G. The promise of record linkage for assessing the uptake of health services in resource constrained settings: a pilot study from South Africa. BMC Med Res Methodol. 2014 May;24(1):71.
21. West RL, Lippman SA, Twine R, Maritze M, Kahn K, Leslie HH. Providers' definitions of quality and barriers to providing quality care: a qualitative study in rural Mpumalanga Province, South Africa. Journal of Global Health Science [Internet]. 2021 Jun 22 [cited 2021 Jul 14];3. Available from: https://doi.org/10.35500/jghs.2021.3.e1
22. Johnston LG, Sabin ML, Prybylski D, Sabin K, McFarland W, Baral S et al. The importance of assessing self-reported HIV status in bio-behavioural surveys. Bull World Health Organ. 2016 Aug 1;94(8):605–12.
23. Mooney AC, Campbell CK, Ratlhagana MJ, Grignon JS, Mazibuko S, Agnew E et al. Beyond Social Desirability Bias: Investigating Inconsistencies in Self-Reported HIV Testing and Treatment Behaviors Among HIV-Positive Adults in North West Province, South Africa. AIDS Behav. 2018 Jul 1;22(7):2368–79.
24. Weiser SD, Heisler M, Leiter K, Percy-de Korte F, Tlou S, DeMonner S, et al. Routine HIV testing in Botswana: a population-based study on attitudes, practices, and human rights concerns. PLoS Med. 2006 Jul;3(7):e261.
25. Turan JM, Bukusi EA, Onono M, Holzemer WL, Miller S, Cohen CR. HIV/AIDS stigma and refusal of HIV testing among pregnant women in rural Kenya: results from the MAMAS Study. AIDS Behav. 2011 Aug;15(6):1111–20.
26. Monitoring PEPFAR. Evaluation, and Reporting Database [Internet]. [cited 2023 Jan 9]. Available from: https://mer.amfar.org/location/South%20Africa/HTS_TST
27. Department of Cooperative Governance and Traditional Affairs, Profile. Ehlanzeni District. 2020 Jul 7 [cited 2023 Jan 9]; Available from: https://www.cogta.gov.za/ddm/wp-content/uploads/2020/07/Take3_Final-Edited-Ehlanzeni-DM_07July2020-FINAL.pdf
28. Little RJA, Rubin DB. Statistical analysis with missing data. NY: Sons; 2002.
29. Jooste S, Mabaso M, Taylor M, North A, Shean Y, Simbayi LC, et al. Geographical variation in HIV testing in South Africa: evidence from the 2017 national household HIV survey. South Afr J HIV Med. 2021 Aug;31(1):1273.
30. Pettifor A, Lippman SA, Kimaru L, Haber N, Mayakayaka Z, Selin A et al. HIV self-testing among young women in rural South Africa: A randomized controlled trial comparing clinic-based HIV testing to the choice of either clinic testing or HIV self-testing with secondary distribution to peers and partners. EClinicalMedicine. 2020 Apr 1;21:100327.
31. Yorlets RR, Lurie MN, Ginsburg C, Hogan JW, Joyce NR, Harawa S et al. Validity of Self-Report for Ascertaining HIV Status Among Circular Migrants and Permanent Residents in South Africa: A Cross-Sectional, Population-Based Analysis. AIDS Behav [Internet]. 2022 Sep 16; Available from: https://doi.org/10.1007/s10461-022-03828-w
32. West RL. Using multiple modes of assessment to measure patient experience in public clinics in rural Mpumalanga, south Africa. Oral presentation presented at: South Africa AIDS Conference; 2019 Jun 12; Durban, South Africa.
33. Chasimpha SJD, Mclean EM, Dube A, McCormack V, dos-Santos-Silva I, Glynn JR. Assessing the validity of and factors that influence accurate self-reporting of HIV status after testing: a population-based study. AIDS. 2020 Jan;30(6):931–41.
34. Rohr JK, Xavier Gómez-Olivé F, Rosenberg M, Manne-Goehler J, Geldsetzer P, Wagner RG, et al. Performance of self-reported HIV status in determining true HIV status among older adults in rural South Africa: a validation study. J Int AIDS Soc. 2017;20(1):21691.