Published in final edited form as:

J Epidemiol Community Health. 2018 February ; 72(2): 168–172. doi:10.1136/jech-2017-209713.

High percentage of undiagnosed HIV cases within a hyperendemic South African community: a population-based study

Alain Vandormael^{1,2,3}, Tulio de Oliveira^{3,4}, Frank Tanser^{1,2,4,5}, Till Bärnighausen^{1,5,6,7}, and Joshua T Herbeck⁸

¹Africa Health Research Institute (AHRI), Durban, South Africa

²School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa

³KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), University of KwaZulu-Natal, South Africa

⁴Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa

⁵Department of Infection and Population Health, Institute of Epidemiology and Health Care, University College London, London, UK

⁶Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Massachusetts, USA

⁷Heidelberg Institute for Public Health, University of Heidelberg, Heidelberg, Germany

⁸International Clinical Research Center, Department of Global Health, University of Washington, Seattle, Washington, USA

Abstract

Background—Undiagnosed HIV infections could undermine efforts to reverse the global AIDS epidemic by 2030. In this study, we estimated the percentage of HIV-positive persons who remain undiagnosed within a hyperendemic South African community.

Methods—The data come from a population-based surveillance system located in the Umkhanyakude district of the northern KwaZulu-Natal province, South Africa. We annually tested 38661 adults for HIV between 2005 and 2016. Using the HIV-positive test results of 12039 (31%) participants, we then back-calculated the incidence of infection and derived the number of undiagnosed cases from this result.

Correspondence to Dr Alain Vandormael, School of Nursing and Public Health, University of KwaZulu-Natal, Durban 4001, South Africa; avandormael@ahri.org.

Contributors AV undertook the statistical analysis and wrote the manuscript. AV and JTH conceived the study. JTH supervised the study, with additional advice, feedback and comments provided by FT, TdO and TB.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Ethics approval for data collection and use was obtained from the biomedical and ethics committee of the University of KwaZulu-Natal (Durban, South Africa).

Provenance and peer review Not commissioned; externally peer reviewed.

Results—The percentage of undiagnosed HIV cases decreased from 29.3% in 2005 to 15.8% in 2011. During this period, however, approximately 50% of the participants refused to test for HIV, which lengthened the average time from infection to diagnosis. Consequently, the percentage of undiagnosed HIV cases reversed direction and steadily increased from 16.1% to 18.9% over the 2012–2016 period.

Conclusions—Results from this hyperendemic South African setting show that the HIV testing rate is low, with long infection times, and an unsatisfactorily high percentage of undiagnosed cases. A high level of repeat HIV testing is needed to minimise the time from infection to diagnosis if the global AIDS epidemic is to be reversed within the next two decades.

INTRODUCTION

Undiagnosed HIV infections could potentially undermine efforts to reverse the AIDS pandemic within the next two decades.¹ These efforts rely on the early initiation of combination antiretroviral therapy (ART), which can prevent the transmission of HIV between infected persons and their uninfected partners.^{2–5} For this reason, the Joint United Nations Programme on HIV/AIDS (UNAIDS) has set an ambitious target to reduce the undiagnosed population size to the 10% level (or below) by the year 2020.⁶ While some mathematical models predict this target can be reached,⁷ real-world data are needed to evaluate progress. Such data could also be used to reveal the limitations of existing HIV testing strategies and inform the future need for ART.^{8–10}

Population-based surveillance systems are a good source of data for estimating the number of undiagnosed HIV cases. These estimates can be derived from a back-calculation approach that uses the number of HIV/AIDS diagnoses recorded for each year of the surveillance period.¹¹ However, several variants of this approach require additional data to improve inference about the distribution of time from infection to diagnosis. For example, the London,¹² Cambridge,¹³ Atlanta,¹⁴ Ottawa/Sydney¹⁵ and Paris¹⁶ methods use either routinely collected CD4+ T cell count data, the clinical status of the patient at diagnosis, or the diagnosis of a prior seroconversion illness. But such clinical and/or laboratory data may be too costly to obtain in most sub-Saharan African settings, even if population-based surveillance systems are currently in place.

The Seattle¹⁷ method is a promising variant that uses data from only the latest HIV-negative and earliest HIV-positive test results. This method is therefore well suited for surveillance systems in which participants are repeatedly tested for HIV. In the case of missing data, a straightforward procedure can be used to impute the HIV-negative test date. The Seattle method also back-calculates the incidence of HIV infection and then derives the number of undiagnosed cases from this result. Using the Seattle method and data from a population-based surveillance system, we estimated the percentage of undiagnosed HIV cases in a hyperendemic South African community between 2005 and 2016.

METHODS

Surveillance system

The Africa Health Research Institute (AHRI) maintains a population-based surveillance system in the Umkhanyakude district of the KwaZulu-Natal province. The surveillance area is 438 km² in size with a population of approximately 90 000 people.¹⁸ Most of the surveillance area is poor and rural, with several informal periurban settlements and a single urban township. Households tend to be scattered across the landscape. Annual HIV testing has been ongoing since January 2004. Trained field-workers visit households every 12 months and identify eligible participants older than 15 years for HIV testing. After obtaining consent, the fieldworkers then extract blood according to the UNAIDS and WHO *Guidelines for Using HIV Testing Technologies in Surveillance*.

Previous work has shown a high burden of HIV in the surveillance area. Between 2004 and 2014, the overall HIV incidence was 3.7 new infections per 100 person-years (95% CI 3.6 to 3.9).¹⁹ By sex and age group, incidence was highest in women aged 20–24 years, at 7.8 new infections per 100 person-years (95% CI 7.2 to 8.5); and highest in men aged 25–29 years, at 4.7 new infections per 100 person-years (95% CI 3.8 to 5.7). The prevalence of HIV steadily increased from approximately 22% in 2005 to 37% in 2016, due, in large part, to the rapid and comprehensive scale-up of ART.^{20,21} ART was made available to patients with CD4+ T cell counts <200 cells/μL through public sector healthcare clinics in 2004 and then to all adult patients in 2011.²²

Seattle method

The Seattle¹⁷ method relies on the participant's latest HIV-negative and earliest HIV-positive test dates. Using these data, the method estimates the population distribution of time from infection to HIV diagnosis (TID). Thereafter, the TID is used to calculate the annual number of undiagnosed HIV cases. In our study, 9510 (79%) of the 12 039 HIV-positive participants did not have an HIV-negative test date. We therefore had to impute a latest HIV-negative test date for these participants.

Imputation of the latest HIV-negative date

The imputation procedure for the latest HIV-negative test date is based on the Seattle method. Using the following notation, let i denote the i th HIV-positive participant for $i = 1, \dots, n$, where $n=12039$. Let l_i be the latest HIV-negative test date and r_i be the earliest HIV-positive test date, and let $x_i = r_i - l_i$, which we call the censored interval. Further, let $s = 1$ January 2005 be the start of the observation period and the date at which undiagnosed time begins to accumulate.

We first defined a maximum censored interval length for all participants with a missing HIV-negative test date. To do this, we used the result of Morgan *et al*,²³ who showed that the time from infection to AIDS was approximately 10 years for 75% of persons living in a rural sub-Saharan African setting. We therefore subtracted 10 years from the earliest HIV-positive test date (r_i) to obtain the imputed HIV-negative test date (l_i), with two exceptions to this rule. First, if this '10 year' imputed HIV-negative date came before the start date (s), then we

made it equal to the start date (ie $I_i = s$ if $I_i < s$). For example, if a 30-year-old participant had an earliest HIV-positive test date on 1 July 2013, then we imputed the HIV-negative date on 1 January 2005. As a sensitivity analysis, we also evaluated 11, 12 and 13-year periods from infection to AIDS. Second, if the imputed HIV-negative date came before the participant's 16th birthdate (b_i), then we made it equal to the 16th birthdate (ie, $I_i = b_i$ if $I_i < b_i$). We use the 16th birthdate to represent the age of first sex and the earliest possible date of HIV infection. This second assumption therefore excludes earlier HIV infection events that could have been acquired vertically or through other transmission routes. As a sensitivity analysis, we also evaluated a 15th birthdate of sexual debut (but could not go lower than this birthdate given our 15-year age eligibility criteria for HIV testing).

Since some participants could potentially have three different imputed dates, we selected the minimum censored interval length for each participant, that is, $x_i = \min([r_i - I_i], [r_i - b_i], [r_i - s])$. For these 9510 participants, we imputed a latest HIV-negative date at the start date for 7171 (75%) participants, at the 16th birthdate for 1228 (13%) participants and at a 10-year date prior to their earliest HIV-positive test date for 1111 (12%) participants.

Estimating the TID

The Seattle method uses the i th participant's censored interval to estimate the population distribution of the TID. Because the exact infection date is rarely observed during periodic testing, the method performs a sensitivity analysis using two different assumptions: (1) the base case (BC), which places an equal probability of infection on every time point within the censored interval; and (2) the upper bound (UB), which assumes that the infection date occurs 1 day after the latest HIV-negative test date. A lower bound assumption is not evaluated here because, from an HIV treatment-as-prevention (TasP) perspective, we are interested in the worst-case scenario, which is the maximum possible number of undiagnosed cases.

Following the notation of Fellows *et al.*¹⁷ let the cumulative hazard of diagnosis be defined as $F(t) = P(TID \leq t)$ and the survival function of the remaining diagnoses as $S(t) = 1 - F(t) = P(TID > t)$. Then $F(t)$ is evaluated under the UB assumption as:

$$\hat{F}_{UB}(t) = \frac{1}{n} \sum_{i=1}^n I(x_i \leq t) \quad (1)$$

where I is an indicator function ($x_i \leq t = 1$) and $\hat{F}_{UB}(t)$ is the empirical cumulative density function of the TID. We note that the UB assumption represents the maximum amount of time that an infected participant can remain undiagnosed. It does not reflect a reason or motivation for the participant to test at any one point in time, as this is determined by the annual testing schedule of the AHRI surveillance system. Under the BC assumption, $F(t)$ is evaluated as:

$$\hat{F}_{BC}(t) = \frac{1}{n} \sum_{i=1}^n \min(1, \frac{t}{x_i}) \quad (2)$$

Equation (2) shows that the BC assumption does not impute a random infection event within the censored interval. Rather, the cumulative probability of infection within the censored interval is 100%—and this probability has a uniform distribution.

The discretised TID density is then obtained for the BC and UB assumptions for each year of the observation period, given as $f_t = F(t+1) - F(t)$, which is the probability that a participant is diagnosed t years after infection.

Calculation of the undiagnosed HIV cases

The Seattle method then uses a Poisson process with rate λ_i to back-calculate the yearly HIV incidence.¹⁷ Once the incidence rate is obtained, the number of undiagnosed cases can be estimated with:

$$E(U_t) = \sum_{i \leq t} \lambda_i (\frac{1}{2} f_{t-i} + \sum_{k < t-i} f_k) \quad (3)$$

For this calculation, each term in the outer sum represents the expected number infected during year t who are diagnosed after year t plus half the expected number of cases diagnosed during year t . To calculate the undiagnosed percentage, we divided the undiagnosed counts, $E(U_t)$, by the total number of HIV-positive participants in our surveillance area for year t under the BC and UB assumptions. For visualisation purposes, we fitted a kernel smoother to the estimates. We performed all analyses in R software V. 3.4.0 using the *HIVBackCalc* package.

RESULTS

Between 2005 and 2016, there were 65473 adults aged 16–55 years who participated in our HIV surveillance system. Of these participants, 38 661 (59.1%) had at least one valid HIV test result and 12 039 (31.1%) had an HIV-positive test result. The median (IQR) age of the study sample at diagnosis was 30 (23–38) years and 73.5% were female, as shown in table 1.

Table 2 shows that approximately 20 000 adults participated in each year of the surveillance programme, with a minimum of 18 527 participants in 2012 and a maximum of 29 824 participants in 2005. The percentage of participants that ever tested for HIV was 30.9% in 2005, which increased to 61.7% in 2010 and then 75.9% in 2016. Refusal rates varied during this time, peaking at 70.2% in 2008, and then declining to 46.5% in 2014 and 41.3% in 2016. Less than 25% of the participants who tested for HIV in our surveillance system were aware of their status. This number peaked at 43.6% in 2012 and then dropped to 35.8% in 2016. The HIV prevalence rose steadily from 21.7% in 2005 to 28.7% in 2010 and to 36.6% in 2016.

Results show that 29.3% (UB=47.1%) of the participants were undiagnosed at the start of the observation period in 2005. Thereafter, the undiagnosed percentage declined to 19.3% (UB=36.3%) in 2007 and to 15.8% (UB=32.8%) in 2011. From 2012 onwards, the percentage of undiagnosed cases increased steadily to 18.9% (UB=37.6%) in 2016, as shown in figure 1.

As a sensitivity analysis, we also evaluated 11, 12 and 13-year periods from infection to AIDS. Results show that longer periods increased the average censored interval length, which consequently increased the number of undiagnosed HIV cases. Using 2016 data as an example, the BC estimate under the 10-year infection period rose from 18.86% (UB=37.6%) to 19.20% (UB=39.35%) under the 11-year period, and then to 21.88% (UB=46.19%) under the 12-year period, as shown in online supplementary table S1. Online supplementary table S2 shows the results for the 15th birthdate assumption under the three infection periods. Lowering the birthdate also increased the average length of the censored interval, leading to a slight increase in the percentage of undiagnosed HIV cases. The estimates we report in table 2, which use a 10-year infection period and a 16th birthdate, can therefore be considered as conservative, since longer infection periods and a lower birthdate lead to a higher percentage of undiagnosed cases within the surveillance area.

DISCUSSION

We have estimated the percentage of undiagnosed cases within a hyperendemic South African community, using data from one of the world's largest HIV surveillance programmes. Our results show that the percentage of undiagnosed cases declined from 29.3% (UB=47.1%) in 2005 to 15.8% (UB=32.8%) in 2011, but then steadily increased back to 18.9% (UB=37.6%) in 2016. This 2016 estimate is far from the UNAIDS target to reduce the undiagnosed population size to the 10% level (or below) by the year 2020.

The increase in the percentage of undiagnosed HIV cases between 2012 and 2016 is due to the high percentage of participants who refused to test or who missed their scheduled test dates. Refusal rates peaked at 70% in 2008 and then declined to 41.3% in 2016, which is an alarmingly high percentage. By the end of 2016, approximately 25% of all participants in our surveillance area had still not tested for HIV, despite ongoing annual testing since 2005. Overall, only 20% of the participants in our surveillance area continually consent to testing, whereas 37% refuse to test and 43% switch between consent and refusal.²⁴ The low HIV testing rate is likely due to a range of social and behavioural factors associated with testing fatigue,²⁴ poverty and inequality,²⁵ social disequilibrium,²⁶ stigma²⁷ and frequent cyclical migration²⁸ within the AHRI surveillance area.

The low testing rate can explain, in large part, the high percentage of undiagnosed cases over the last 5 years of the observation period. As previous research has shown, missed test dates will lengthen the time between the latest HIV-negative and earliest HIV-positive test dates,²⁹ otherwise known as the censored interval. In our surveillance cohort, the average censored interval length increased from 1.1 years in 2005 to 4.3 years in 2016. Thus, the early decline in the percentage of undiagnosed cases over the 2005–2011 period was likely offset by an

increase in the amount of undiagnosed time, which led to a greater percentage of undiagnosed cases from 2012 onwards.

One possible limitation of this study is that a substantial percentage (79%) of the 12 039 participants did not have an HIV-negative test result, which is a common problem for large HIV surveillance systems. We therefore had to infer a latest HIV-negative date for these participants: either 10 years prior to the earliest HIV-positive test date, at the start of the study period, or at the 16th birthdate—whichever gave the shortest censored interval. By selecting the shortest interval, it is likely that we reduced the average amount of undiagnosed time, and therefore underestimated the percentage of undiagnosed HIV cases. Our results should therefore be considered as conservative, as demonstrated by our sensitivity analyses using 11, 12 and 13-year periods prior to the earliest HIV-positive test date.

Our analyses assumed that the HIV infection event occurred no earlier than the participant's 16th birthdate—the age of sexual debut. Although the heterosexual transmission of HIV is the predominant route of infection in sub-Saharan Africa,³⁰ a small but unknown proportion of participants in our study area could have acquired HIV vertically (or via other transmission routes) before their 16th birthdate. Nevertheless, the aim of this study was to quantify the percentage of undiagnosed HIV cases using data from sexually active participants, since TasP efforts are currently focused on preventing the onward transmission of HIV between serodiscordant partners.⁶ Further, our sensitivity analysis showed that lowering the age of sexual debut by 1 year only increased the percentage of undiagnosed HIV cases, thus confirming the main conclusion of this study. Under the 16th birthdate assumption, our estimates can therefore be considered as conservative.

We also made the additional assumption that the number of undiagnosed cases could be reliably derived from a back-calculated HIV incidence rate. Lodwick *et al*¹ provide a more detailed discussion of this assumption elsewhere. Nevertheless, our estimates correspond closely with the Agence Nationale de Recherche sur le Sida (ANRS) TasP trial, which was undertaken in 22 communities adjacent to our surveillance area between March 2012 and April 2016.³¹ Investigators report that between 15% and 26% of the 14 916 participants did not receive an HIV diagnosis, a range which includes our 16%–21% estimates over the 2005–2016 observation period. Results from both of these studies, which use real-world data, are higher than the 10% undiagnosed size for 2016, as projected by a previously published simulation model.⁷

We show in a hyperendemic South African setting that the HIV testing rate is low, with long infection times, and an unsatisfactorily high percentage of undiagnosed cases. This scenario is problematic from a public health perspective: undiagnosed and untreated individuals will remain infectious for longer periods and have a higher probability of transmitting HIV to their uninfected sexual partners.^{32,33} A high level of HIV testing is needed to minimise the time to diagnosis and reduce the number of undiagnosed cases if the global AIDS epidemic is to be reversed by 2030.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding This work was supported by a South African Medical Research Council (SA MRC) Flagship grant (MRC-RFA-UFSP-01-2013/UKZN HIVEPI). Funding for the Africa Health Research Institute's Demographic Surveillance Information System and Population-based HIV Survey was received from the Wellcome Trust. FT was supported by two National Institute of Health (NIH) grants (R01HD084233 and R01AI124389) as well as a UK Academy of Medical Sciences Newton Advanced Fellowship (NA150161). JTH was supported by two NIH grants (R01AI108490 and R01AI127232). TB was supported by the Alexander von Humboldt Foundation through the endowed Alexander von Humboldt Professorship funded by the German Federal Ministry of Education and Research, as well as by the Wellcome Trust, the European Commission, the Clinton Health Access Initiative, and by four NIH grants (R01-HD084233, R01-AI124389, D43-TW009775, R01-AI112339).

REFERENCES

1. UNAIDS. Fast-Track - Ending the AIDS epidemic by 2030. Geneva: Joint United Nations Programme on HIV/AIDS, 2014.
2. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med* 2011;365:493–505. [PubMed: 21767103]
3. Baeten JM, Donnell D, Ndase P, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *N Engl J Med* 2012;367:399–410. [PubMed: 22784037]
4. Rodger AJ, Cambiano V, Bruun T, et al. Sexual activity without condoms and risk of HIV transmission in serodifferent couples when the HIV-Positive partner is using suppressive antiretroviral therapy. *JAMA* 2016;316:171–81. [PubMed: 27404185]
5. Lundgren JD, Babiker AG, Gordin F, et al. Initiation of antiretroviral therapy in early asymptomatic HIV infection. *N Engl J Med* 2015;373:795–807. [PubMed: 26192873]
6. UNAIDS. 90-90-90: an ambitious treatment target to help end the AIDS epidemic. Geneva: Joint United Nations Programme on HIV/AIDS, 2014.
7. Johnson LF, Rehle TM, Jooste S, et al. Rates of HIV testing and diagnosis in South Africa: successes and challenges. *AIDS* 2015;29:1401–9. [PubMed: 26091299]
8. Fisher M Late diagnosis of HIV infection: major consequences and missed opportunities. *Curr Opin Infect Dis* 2008;21:1–3. [PubMed: 18192778]
9. Antinori A, Coenen T, Costagliola D, et al. Late presentation of HIV infection: a consensus definition. *HIV Med* 2011;12:61–4. [PubMed: 20561080]
10. Mammone A, Pezzotti P, Regine V et al. How many people are living with undiagnosed HIV infection? An estimate for Italy, based on surveillance data. *AIDS* 2016;30:1131–6. [PubMed: 26807973]
11. Lodwick R, Alioum A, Archibald C, et al. HIV in hiding: methods and data requirements for the estimation of the number of people living with undiagnosed HIV. *AIDS* 2011;25:1017–23. [PubMed: 21422986]
12. Lodwick R, Sabin C, Phillips A. Estimation of the number of undiagnosed HIV-positive people within a region based on surveillance of simultaneous HIV/AIDS diagnoses. *HIV Medicine* 2009;10:193.
13. Sweeting MJ, De Angelis D, Aalen OO. Bayesian back-calculation using a multi-state model with application to HIV. *Stat Med* 2005;24:3991–4007. [PubMed: 16320278]
14. Hall HI, Song R, Rhodes P, et al. Estimation of HIV incidence in the United States. *JAMA* 2008;300:520–9. [PubMed: 18677024]
15. Wand H, Yan P, Wilson D, et al. Increasing HIV transmission through male homosexual and heterosexual contact in Australia: results from an extended back-projection approach. *HIV Med* 2010;11:395–403. [PubMed: 20136660]

16. Ndawinz JD, Costagliola D, Supervie V. New method for estimating HIV incidence and time from infection to diagnosis using HIV surveillance data: results for France. *AIDS* 2011;25:1905–13. [PubMed: 21811147]
17. Fellows IE, Morris M, Birnbaum JK, et al. A new method for estimating the number of undiagnosed HIV infected based on HIV testing history, with an application to men who have sex with men in Seattle/King County, WA. *PLoS One* 2015;10:e0129551. [PubMed: 26196132]
18. Tanser F, Hosegood V, Bärnighausen T, et al. Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. *Int J Epidemiol* 2008;37:956–62. [PubMed: 17998242]
19. Vandormael A, Newell ML, Bärnighausen T, et al. Use of antiretroviral therapy in households and risk of HIV acquisition in rural KwaZulu-Natal, South Africa, 2004–12: a prospective cohort study. *Lancet Glob Health* 2014;2:e209–e215. [PubMed: 24782953]
20. Bor J, Herbst AJ, Newell ML, et al. Increases in adult life expectancy in rural South Africa: valuing the scale-up of HIV treatment. *Science* 2013;339:961–5. [PubMed: 23430655]
21. Zaidi J, Grapsa E, Tanser F, et al. Dramatic increases in HIV prevalence after scale-up of antiretroviral treatment: a longitudinal population-based HIV surveillance study in rural kwazulu-natal. *AIDS* 2013;27:2301. [PubMed: 23669155]
22. Houlihan CF, Bland RM, Mutevedzi PC, et al. Cohort profile: Hlabisa HIV treatment and care programme. *Int J Epidemiol* 2011;40:318–26. [PubMed: 20154009]
23. Morgan D, Mahe C, Mayanja B, et al. HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? *AIDS* 2002;16:597–603. [PubMed: 11873003]
24. Larmarange J, Mossong J, Bärnighausen T, et al. Participation dynamics in population-based longitudinal HIV surveillance in rural South Africa. *PLoS One* 2015;10:e0123345. [PubMed: 25875851]
25. Thurlow J, Gow J, George G. HIV/AIDS, growth and poverty in KwaZulu-Natal and South Africa: an integrated survey, demographic and economy-wide analysis. *J Int AIDS Soc* 2009;12:18. [PubMed: 19758444]
26. Tomita A, Vandormael AM, Bärnighausen T, et al. Social disequilibrium and the risk of HIV acquisition: a multilevel study in Rural KwaZulu-Natal Province, South Africa. *J Acquir Immune Defic Syndr* 2017;75:164–74. [PubMed: 28291049]
27. Tanser F, Bärnighausen T, Vandormael A, et al. HIV treatment cascade in migrants and mobile populations. *Curr Opin HIV AIDS* 2015;10:430–8. [PubMed: 26352396]
28. Dobra A, Bärnighausen T, Vandormael A, et al. Space-time migration patterns and risk of HIV acquisition in rural South Africa: a population-based cohort study. *AIDS* 2017;31:137–45. [PubMed: 27755099]
29. Vandormael A, Dobra A, Bärnighausen T, et al. Incidence rate estimation, periodic testing and the limitations of the mid-point imputation approach. *Int J Epidemiol*. In Press. 2017;2017.
30. Unaid U The gap report. Geneva, Switzerland 2014.
31. Iwuji C, Orne-Gliemann J, Balestre E, et al. The impact of universal test and treat on HIV incidence in a rural South African population: ANRS 12249 TasP trial, 2012–(95%2016. *AIDS* 2016.
32. Hughes JP, Baeten JM, Lingappa JR, et al. Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples. *J Infect Dis* 2012;205:358–65. [PubMed: 22241800]
33. Wawer MJ, Gray RH, Sewankambo NK, et al. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J Infect Dis* 2005;191:1403–9. [PubMed: 15809897]

What is already known on this subject

Current treatment-as-prevention strategies aim to reduce the size of the undiagnosed HIV population to the 10% level (or below) by the year 2020. Some mathematical models predict this target can be reached; however, real-world data are critically needed to evaluate progress.

What this study adds

Using data from a population-based surveillance system, we calculated the percentage of undiagnosed HIV cases in a hyperendemic South African setting. Our results show that the percentage of undiagnosed cases was 18.9% in 2016, with an upper bound of 37.6%—much higher than the 10% target set by the Joint United Nations Programme on HIV/AIDS. A high level of repeat HIV testing is needed to minimise the time from infection to diagnosis.

Figure 1.

The percentage of undiagnosed HIV cases over the 2005–2016 observation period. The dashed line shows the base case (BC) estimates (smoothed) and the solid line shows the upper bound (UB) estimates (smoothed). The BC assumption places an equal probability of infection on every time point between the latest HIV-negative and earliest HIV-positive test dates, and the UB assumes that the infection event occurs 1 day after the latest HIV-negative test date.

Table 1

Testing characteristics of the study population by sex, 2005–2016. The censored interval is the time between the latest HIV-negative and the earliest HIV-positive test date

Testing characteristics	All participants		Female		Male	
Total tested, n (%)	65 473		35 160	53.7	30 313	46.3
Valid HIV test result, n (%)	38 661	59.1	22 665	64.5	15 996	52.8
HIV-positive result, n (%)	12 039	31.1	8852	39.1	3187	19.9
Has HIV-negative result, n (%)	2529	21.0	1948	22.0	581	18.2
Censored interval length (years), median (IQR)	3.26	(1.06–6.43)	3.15	(1.04–6.28)	3.59	(1.19–6.80)
Censored interval length (years), mean (SD)	4.12	3.24	4.06	3.22	4.29	3.28
Age at diagnosis, median (IQR)	30	(23–38)	29	(23–37)	32	(26–40)
HIV-positive women, n (%)	8862	73.5				

Table 2

Number of participants who tested each year, the percentage who ever tested for HIV, the percentage who refused an HIV test, the percentage who were aware of their HIV status, the HIV prevalence and the percentage of undiagnosed HIV cases between 2005 and 2016. The prevalence is calculated as the percentage of HIV-positive cases among the general population with a valid HIV test result

Year	Total tested	Ever tested	Refused	Status aware	HIV prevalence		HIV undiagnosed*	
	n	%	%	%	%	(95%CI)	BC (%)	UB (%)
2005	29 824	30.93	44.57		21.68	(20.83 to 22.53)	29.28	47.11
2006	21 817	49.03	61.79	24.87	21.40	(20.50 to 22.30)	20.77	37.42
2007	21 759	52.94	65.68	24.93	23.08	(22.13 to 24.03)	19.30	36.26
2008	24 971	53.91	70.23	23.93	23.65	(22.69 to 24.61)	18.25	35.30
2009	22 024	57.26	68.69	32.18	25.66	(24.63 to 26.69)	17.99	35.22
2010	22 227	61.67	43.29	33.38	28.66	(27.70 to 29.62)	17.56	34.24
2011	20 766	65.02	39.18	39.93	28.07	(27.09 to 29.06)	15.77	32.79
2012	18 527	64.01	35.00	43.57	30.06	(28.89 to 31.22)	16.13	33.89
2013	20 326	66.66	40.32	35.57	32.17	(31.11 to 33.24)	17.32	35.45
2014	20 064	67.93	46.51	30.63	34.82	(33.70 to 35.94)	18.10	36.72
2015	22 024	72.72	40.31	34.48	34.75	(33.81 to 35.69)	19.15	37.82
2016	22 576	75.94	41.32	35.75	36.59	(35.69 to 37.50)	18.86	37.60

* Because the exact infection date is rarely observed during periodic testing, the Seattle method calculates the undiagnosed percentage using two assumptions: (1) the base case (BC), which places an equal probability of infection on every time point between the latest HIV-negative and earliest HIV-positive test dates; and (2) the upper bound (UB), which assumes that the infection event occurs 1 day after the latest HIV-negative test date.