

# HIV Incidence in Rural South Africa: Comparison of Estimates from Longitudinal Surveillance and Cross-Sectional cBED Assay Testing

Till Bärnighausen<sup>1\*</sup>, Claudia Wallrauch<sup>1</sup>, Alex Welte<sup>2,3</sup>, Thomas A. McWalter<sup>2</sup>, Nhlanhla Mbizana<sup>1</sup>, Johannes Viljoen<sup>1</sup>, Natalie Graham<sup>1</sup>, Frank Tanser<sup>1</sup>, Adrian Puren<sup>4</sup>, Marie-Louise Newell<sup>1,5</sup>

**1** Africa Centre for Health & Population Studies, University of KwaZulu-Natal, Durban, South Africa, **2** School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa, **3** DST/NRF Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa, **4** National Institute of Communicable Diseases, Johannesburg, South Africa, **5** Centre for Paediatric Epidemiology and Biostatistics, Institute of Child Health, University College of London, London, United Kingdom

## Abstract

**Background:** The BED IgG-Capture Enzyme Immunoassay (cBED assay), a test of recent HIV infection, has been used to estimate HIV incidence in cross-sectional HIV surveys. However, there has been concern that the assay overestimates HIV incidence to an unknown extent because it falsely classifies some individuals with non-recent HIV infections as recently infected. We used data from a longitudinal HIV surveillance in rural South Africa to measure the fraction of people with non-recent HIV infection who are falsely classified as recently HIV-infected by the cBED assay (the long-term false-positive ratio (FPR)) and compared cBED assay-based HIV incidence estimates to longitudinally measured HIV incidence.

**Methodology/Principal Findings:** We measured the long-term FPR in individuals with two positive HIV tests (in the HIV surveillance, 2003–2006) more than 306 days apart (sample size  $n = 1,065$ ). We implemented four different formulae to calculate HIV incidence using cBED assay testing ( $n = 11,755$ ) and obtained confidence intervals (CIs) by directly calculating the central 95<sup>th</sup> percentile of incidence values. We observed 4,869 individuals over 7,685 person-years for longitudinal HIV incidence estimation. The long-term FPR was 0.0169 (95% CI 0.0100–0.0266). Using this FPR, the cross-sectional cBED-based HIV incidence estimates (per 100 people per year) varied between 3.03 (95% CI 2.44–3.63) and 3.19 (95% CI 2.57–3.82), depending on the incidence formula. Using a long-term FPR of 0.0560 based on previous studies, HIV incidence estimates varied between 0.65 (95% CI 0.00–1.32) and 0.71 (95% CI 0.00–1.43). The longitudinally measured HIV incidence was 3.09 per 100 people per year (95% CI 2.69–3.52), after adjustment to the sex-age distribution of the sample used in cBED assay-based estimation.

**Conclusions/Significance:** In a rural community in South Africa with high HIV prevalence, the long-term FPR of the cBED assay is substantially lower than previous estimates. The cBED assay performs well in HIV incidence estimation if the locally measured long-term FPR is used, but significantly underestimates incidence when a FPR estimate based on previous studies in other settings is used.

**Citation:** Bärnighausen T, Wallrauch C, Welte A, McWalter TA, Mbizana N, et al. (2008) HIV Incidence in Rural South Africa: Comparison of Estimates from Longitudinal Surveillance and Cross-Sectional cBED Assay Testing. PLoS ONE 3(11): e3640. doi:10.1371/journal.pone.0003640

**Editor:** Landon Myer, University of Cape Town, South Africa

**Received:** July 12, 2008; **Accepted:** October 17, 2008; **Published:** November 4, 2008

**Copyright:** © 2008 Bärnighausen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by grants from the Wellcome Trust through grants to the Africa Centre for Health and Population Studies, University of KwaZulu-Natal, South Africa (GR065377/Z/01/B), and the Africa Centre's Demographic Surveillance Information System (GR065377/Z/01/H), which funds the population-based HIV surveillance. In addition, this publication was supported by Cooperative Agreement Number 5U62PS022901 from Centers for Disease Control and Prevention (CDC). Till Bärnighausen and Frank Tanser were supported by grant 1R01-HD058482-01 from the National Institute of Child Health and Human Development (NICHD). The contents are solely the responsibility of the authors and do not necessarily represent the official views of the any of the funding organizations. The funding organizations had no role in the design and conduct of the study, in the collection, analysis, and interpretation of the data, or in the preparation, review or approval of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: tbarnighausen@afriacentre.ac.za

## Introduction

To understand the dynamics of the HIV epidemic and to target and evaluate interventions to prevent HIV infection, estimates of HIV incidence at the population level are of prime importance. HIV incidence estimates can be obtained through repeated HIV testing of individuals in longitudinal surveillances. Such surveillances, however, are difficult to establish and expensive to maintain. Longitudinal data on HIV status are thus rarely

available [1]. Alternatively, HIV incidence can be estimated from changes in HIV prevalence over time. The validity of these estimates, however, depends on assumptions about survival time distributions among HIV-positive and -negative individuals, which are commonly quite uncertain [2,3]. Finally, HIV incidence can be measured in a single cross-sectional survey using laboratory tests which distinguish recent from non-recent HIV infections, reducing the need for both longitudinal and repeated cross-sectional measurement in order to estimate HIV incidence [1].

In recent years, a number of large-scale cross-sectional HIV serosurveys have been conducted. For instance, between 2001 and 2008, 20 demographic health surveys (DHS) in developing countries have included nationally representative HIV serosurveys [4]. A valid and affordable laboratory procedure to distinguish between recent and non-recent infections would allow estimation of HIV incidence in these cross-sectional surveys. One serological method to differentiate recent from non-recent HIV infections uses the BED IgG-Capture Enzyme Immunoassay (cBED assay), which measures the proportion of HIV-1-specific IgG out of total IgG. This proportion should increase with time after HIV seroconversion [5]. Seropositive individuals who test below a certain threshold of this proportion (the BED threshold) are classified as recently infected, while those testing above the BED threshold are classified as non-recently infected [5]. The time period following seroconversion after which infections are no longer considered to be recent (the so-called window period of the cBED assay) is usually estimated at approximately half a year [5,6,7].

The cBED assay has been used to estimate HIV incidence in many countries, including in Ethiopia [8], Rwanda [9], South Africa [10,11], Uganda [12], Zambia [9], Zimbabwe [7], China [13,14], and the United States [15,16]. However, there has been concern that the cBED assay-based methods overestimate HIV incidence to an unknown extent because some non-recent infections are classified as recent [17]. In some individuals (so-called non-progressors) the proportion of HIV-1-specific IgG never rises above the recency threshold, and in other individuals (so-called regressors) who have been HIV-infected for a long time, the proportion may fall below the threshold after having previously progressed above it. Regression to levels below threshold can occur for a number of biological reasons that decrease HIV-1-specific IgG relative to total IgG, including viral suppression and immune reconstitution on antiretroviral treatment (ART), concurrent infections, and late-stage HIV disease [17]. It is in principle possible to account for non-recently HIV infected individuals who are misclassified as recently infected, but the HIV incidence estimates will depend on the estimate of a long-term false-positive ratio (FPR) [6,7,18]. All current methods for this correction effectively assume that by some finite time after HIV infection (the maximum BED progression time) all individuals, with the exception of non-progressors, will have progressed to the BED threshold [18]. From previous empirical observations, it is known that the maximum BED progression time is of the order of one year [6,7]. Thus, the fraction of all people who have been HIV-infected at least as long as the maximum BED progression time who are below the BED threshold is the long-term FPR.

We use data from a large population-based longitudinal HIV surveillance to measure the long-term FPR in a rural African community with high HIV prevalence [19] and HIV incidence [20], and then compare HIV incidence estimates based on the cBED assay to estimates based on longitudinal HIV surveillance.

## Methods

### Setting

We used dried blood spot (DBS) specimens which were collected in the longitudinal population-based HIV surveillance conducted by the Africa Centre for Health and Population Studies (Africa Centre), University of KwaZulu-Natal [21]. The HIV surveillance area is located near the market town of Mtubatuba in the Umkhanyakude district of KwaZulu-Natal. The area is 438 square kilometers in size; it has a population of approximately 85,000 almost exclusively Zulu-speaking people who are members of

about 11,000 households [22]. In 2004, the overall HIV prevalence among residents in the surveillance area was 27% in women (15 to 49 years of age) and 14% in men (15 to 54 years of age) [19]. The surveillance methods have been described elsewhere [20,23]. Ethics permission for the HIV surveillance at the Africa Centre was obtained from the Research Ethics Committee at the College of Health Sciences, University of KwaZulu-Natal. All participants in the study provided written informed consent for the analysis of their samples.

### Samples

All women aged 15–49 years and all men aged 15–54 years who were resident in the surveillance area at the time of visit of an HIV surveillance fieldworker were eligible for HIV testing. Different samples were used for the different analyses conducted for this article. The samples for estimation of the long-term FPR consisted of cBED assay results for blood specimens contributed by individuals who tested HIV positive in the surveillance in the time period from June 2003 through June 2006. In order to be included in the sample, the specimens had to meet the following criteria. First, they were follow-up specimens from individuals who had previously tested HIV-positive in the surveillance. Second, the time period between the first positive HIV test and the follow-up specimen exceeded the maximum BED progression time. Third, the specimen was the earliest follow-up specimen that met the second criterion. Our count of long-term false-positive individuals included all individuals who were classified as recently HIV-infected and had been infected for longer than the maximum BED progression time, i.e. it included both non-progressors and regressors.

For the further cBED assay analyses we used a maximum BED progression time of 306 days (sample size  $n = 1,065$ ) as baseline assumption. In order to assess the sensitivity of the long-term FPR to the assumed maximum BED progression time, we varied progression time length from 250 to 400 days in daily intervals. Table 1 shows sample size and the number of individuals who were falsely identified as recently HIV-infected for the BED progression times when the long-term FPR reaches its maximum and minimum and for all progression times in ten-day intervals from 250 to 400 days.

For the HIV incidence estimation based on longitudinal HIV status information, we included all individuals who tested at least twice for HIV in the period from June 2003 through June 2006 and whose first HIV test in this period was negative (4,869 individuals observed over 7,685 person-years). As in previous studies of HIV incidence based on data from longitudinal HIV surveillances [24,25,26,27,28], for the purpose of estimating exposure time, we used the mid-date between the last available negative HIV test and the first available positive HIV test as an estimate of the date of seroconversion. In addition, in order to test the robustness of the longitudinally measured HIV incidence estimates to changes in the assumption about seroconversion dates, we re-estimated HIV incidence using the most extreme assumptions about the seroconversion date that are possible given the interval-censored information on seroconversion dates. At the one extreme, we assumed that all individuals in the longitudinal sample who seroconverted did so on the day immediately after the day of their last HIV-negative test. At the other extreme, we assumed that all individuals who seroconverted did so on the day of their first HIV-positive test. Under changes in the assumption of date of seroconversion, these two extremes yield maximum and minimum estimates of longitudinally measured incidence.

For the cross-sectional cBED-based HIV incidence estimation, we used the first available HIV test for all individuals tested in the

**Table 1.** Long-term FPR.

Maximum BED progression time (in days)	Sample size (individuals)	Number of individuals with false-positive cBED assay results (individuals)	Long-term FPR ( $\epsilon_2$ )	
			Mean	95% CI
250	1100	18	0.0164	0.0097–0.0257
260	1094	18	0.0165	0.0098–0.0259
270	1090	18	0.0165	0.0098–0.0260
280	1083	18	0.0166	0.0099–0.0261
290	1081	18	0.0167	0.0099–0.0262
300	1070	18	0.0168	0.0100–0.0265
<b>306</b>	<b>1065</b>	18	<b>0.0169</b>	<b>0.0100–0.0266</b>
310	1056	18	0.0170	0.0101–0.0268
320	1043	18	0.0173	0.0103–0.0271
330	1035	18	0.0174	0.0103–0.0273
340	1017	18	0.0177	0.0105–0.0278
350	991	17	0.0172	0.0100–0.0273
<b>360</b>	<b>936</b>	<b>17</b>	<b>0.0182</b>	<b>0.0106–0.0289</b>
370	818	14	0.0171	0.0094–0.0285
<b>374</b>	<b>789</b>	<b>14</b>	<b>0.0177</b>	<b>0.0097–0.0296</b>
380	773	14	0.0181	0.0099–0.0302
390	755	14	0.0185	0.0102–0.0309
400	737	14	0.0190	0.0104–0.0317

FPR = false-positive ratio, CI = confidence interval. Row in bold font shows FPR at twice the window period of 153, 180, and 187 days, respectively.  
doi:10.1371/journal.pone.0003640.t001

time period January 2005 through June 2006 ( $n = 11,755$ ), i.e. the period in which all second HIV tests of the people included in the longitudinal HIV incidence analysis took place. Thus, all 4,869 individuals in the longitudinal sample are also included in the sample for the cBED assay-based analysis.

### Laboratory procedures

HIV status was determined by antibody testing with a broad-based HIV-1/HIV-2 enzyme-linked immunosorbent assay (ELISA; Vironostika, Organon Teknika, Boxtel, the Netherlands) followed by a confirmatory ELISA (GAC-ELISA; Abbott, Abbott Park, Illinois, USA) [23]. If HIV-positive status was confirmed, we used another spot from the same filter paper as used for the initial test in order to conduct the cBED assay (cEIA; Calypte<sup>®</sup> HIV-1 BED Incidence EIA, Calypte Biomedical Corporation, Maryland, USA). HIV-specific IgG were detected by the BED-biotin peptide, followed by a colour reaction with streptavidin-peroxidase. The optical density values were normalized in every run using a calibrator (normalized OD (OD<sub>n</sub>) = mean specimen OD/mean calibrator OD). Specimens with OD<sub>n</sub> less than or equal to 1.2 during an initial cBED screening test were confirmed by further cBED testing of the sample in triplicate. We took the median value of the three confirmatory test results as the final OD<sub>n</sub> value. As specified by the manufacturer, an HIV-1-positive specimen for which the cBED assay gave a final OD<sub>n</sub> of less than or equal to 0.8 was considered to be a specimen of recent HIV-1 infection. Otherwise, the specimen was classified as a non-recent infection [5].

### Statistical analysis

Different formulae that use information obtained from the cBED assay have been proposed to estimate HIV incidence from cross-sectional surveys. These formulae provide incidence esti-

mates expressed either as a rate,  $\hat{I}_r$  (expressed, for instance, in number of new HIV infections per 100 person-years) [18] or as the probability that in a given year a person will acquire HIV, i.e. an incidence proportion,  $\hat{I}_p$  (expressed, for instance, in number of new HIV infections per 100 people per year) [6,7]. Some of us have previously derived a formula from first principles to estimate HIV incidence based on the cBED assay [18], and have commented on the assumptions made in different formulae [29,30]. Here, we implemented four different formulae found in the literature. The formula for HIV incidence derived by McDougal and colleagues (McDougal formula) [6] is

$$\hat{I}_p = \frac{fR}{fR + \overline{w}N},$$

where  $R$  is the number of people who were classified as recently HIV-infected by the cBED assay and  $N$  is the number of individuals who tested HIV-negative. The mean window period of the cBED assay,  $\overline{w}$ , is “the mean period of time from initial seroconversion to reaching an OD<sub>n</sub> of 0.800” expressed in years in people who progress above the BED threshold [6]. The “adjustment factor”

$$f = [(R/P) - \epsilon_2] / [(R/P)(\sigma + \epsilon_1 - 2\epsilon_2)]$$

takes into account that the cBED assay does not have perfect specificity or sensitivity,  $P$  is the total number of people who tested HIV-positive,  $\sigma$  is the sensitivity of the cBED assay,  $\epsilon_1$  is the short-term FPR (i.e. over the period  $[\overline{w}, 2\overline{w}]$ ), and  $\epsilon_2$  is the long-term FPR (i.e. over all times  $> 2\overline{w}$ ). Note that the short- and long-term specificities,  $\rho_1$  and  $\rho_2$ , are related to the FPRs by  $\rho_1 = 1 - \epsilon_1$  and  $\rho_2 = 1 - \epsilon_2$ , respectively. The formula of Hargrove and colleagues

(Hargrove formula) [7] is

$$\hat{I}_p = \frac{R - \varepsilon_2 P}{R + \bar{w}N - \varepsilon_2(P + N)},$$

while the formula derived by McWalter and Welte (McWalter/Welte formula) [18] is

$$\hat{I}_r = \frac{R - \left(\frac{\varepsilon_2}{1 - \varepsilon_2}\right)(P - R)}{\bar{w}N}.$$

In addition, we implemented a simplified version of the McDougal formula. The adjustment factor used in the formula can be simplified to

$$f = [(R/P) - \varepsilon_2] / [(R/P)(1 - \varepsilon_2)]$$

using the identity

$$\sigma + \varepsilon_1 - \varepsilon_2 = 1,$$

which requires no more assumptions than are used by McDougal and colleagues [29,30].

Note that in order to implement any of the above four formulae, estimates of the long-term FPR  $\varepsilon_2$  and the window period  $\bar{w}$  are required. For our baseline estimation, we use an  $\bar{w}$  of 153 days, i.e. the window period that is recommended by the manufacturer of the commercially available cBED assay. Most previous studies reporting HIV incidence based on the cBED assay have used window periods between 150 and 160 days [6,9,12,13,14,15,16,31,32,33,34,35]. A few studies have used a window period of 180 days [8,10,11], and a recent study from Zimbabwe calibrated a window period of 187 days in postpartum mothers enrolled in a Vitamin-A intervention trial [7]. In order to test whether our results are robust to changes in the window period estimate, we repeated our analyses with window periods of 180 and 187 days. The Hargrove and McDougal formulae require that the maximum BED progression time is twice the window period. The estimate of the long-term FPR thus depends on the choice of the window period (see Table 1).

Note also that the Hargrove, McWalter/Welte and simplified McDougal formulae do not require estimates of  $\sigma$  and  $\varepsilon_1$ , which – unlike  $\varepsilon_2$  – cannot be calibrated from longitudinal data if the intervals between the last negative and the first positive HIV test in seroconverters are of the order of one year [30]. The mean period of follow-up among seroconverters in our study was 1.4 years; we thus used estimates of  $\sigma$  (0.7680) and  $\varepsilon_1$  (0.2770) from another study in order to implement the McDougal formula [6] (compare also [10]).

The McWalter/Welte formula expresses HIV incidence as a rate, i.e. as the number of HIV seroconversions per person-time at risk, while all other formulae express HIV incidence as an incidence proportion, i.e. the number of HIV seroconversions within a specified time period divided by the size of the population initially at risk. In order to directly compare all HIV incidence estimates in our study, we expressed the estimates based on the McWalter/Welte formula and the longitudinally measured HIV incidence both as rates (per 100 person-years) and as incidence proportions (per 100 people per year). We translated the rate estimates into proportions, assuming that the incidence rate,  $\hat{I}_r$ , is

constant over time  $T$ , by using the relationship

$$\hat{I}_p = 1 - e^{-\hat{I}_r T}.$$

The authors of the four different formulae do not use equivalent methods for the calculation of confidence intervals (CIs). Thus, uncertainty analysis on the incidence estimates was performed as follows. Any observed proportion of HIV-negative, cBED-recent and cBED-non-recent individuals is an unbiased estimate of the underlying population proportions. Given an observed occurrence of the population proportions and the sample size, all attainable draws of the three counts can be enumerated and assigned their respective trinomial probability. Hence an exact cumulative probability distribution of attainable values of the incidence estimator can be computed. For each incidence estimate, we quote the estimator evaluated at the observed counts (the maximum likelihood estimate) and a confidence interval expressed as the central 95<sup>th</sup> percentile.

To control for differences in the sex-age composition between the sample used in the longitudinal HIV incidence estimation and the sample used in the cBED assay-based estimation, we weighted the sex- and five-year age group-specific longitudinal mean incidence rates by the proportions of individuals in each of the sex-age groups in the sample used for the cBED assay-based estimation

$$\hat{I}_{rs} = \sum_i w_{si} \hat{I}_{ri},$$

where  $\hat{I}_{rs}$  is the sex-age adjusted mean incidence rate,  $w_{si}$  are the proportions of individuals in each sex-age group in the cBED assay sample, and  $\hat{I}_{ri}$  are the sex-age specific mean incidence rates. We estimated the variance of  $\hat{I}_{rs}$ ,  $\text{var}(\hat{I}_{rs})$ , as

$$\text{var}(\hat{I}_{rs}) = \sum_i w_{si}^2 \frac{\hat{I}_{ri}^2}{\hat{C}_i}$$

assuming that the number of HIV incident cases,  $\hat{C}_i$ , is Poisson distributed [36]. We calculated the 95% confidence limits for  $\hat{I}_{rs}$  using the method based on gamma distributions described in Anderson and Rosenberg [37].

## Results

### Long-term FPR

Counting the number of DBS specimens classified as recently HIV-infected by the cBED assay in the sample of all individuals who had a previous positive HIV test more than 306 days before the date of the cBED assay-tested specimen, we obtained a long-term FPR of 0.0169 (95% CI 0.0100–0.0266). When we varied the length of the maximum BED progression time from 250 to 400 days (in daily intervals), we found that the estimate of the long-term FPR did not change significantly over the time interval, with minimum and maximum long-term FPRs of 0.0164 (95% CI 0.0097–0.0257) and 0.0190 (95% CI 0.0104–0.0317), respectively (Table 1).

### Incidence comparison

Of the 4,869 individuals included in the sample for longitudinal HIV incidence measurement, 224 people seroconverted in 7,685 person-years. Assuming that seroconversion occurred at the mid-date between the last available negative HIV test and the first

available positive HIV test, longitudinally measured crude HIV incidence was 2.87 per 100 people per year (95% CI 2.53–3.27) (Table 2). Longitudinally measured HIV incidence increased to 3.09 per 100 people per year (95% CI 2.69–3.52), when we adjusted it to the age-sex distribution of the sample for the cBED assay-based incidence estimate.

Of the 11,755 individuals included in the sample for the cBED assay-based HIV incidence measurement, 9,236 tested HIV-negative and 2,519 HIV-positive. Of the individuals who tested HIV-positive, 165 were classified in cBED assay testing as recently HIV-infected and the remainder as non-recently infected. For given  $\epsilon_2$  and  $\bar{w}$ , the four different formulae to calculate HIV incidence from cBED assay measurement produced very similar results. Using the baseline estimate for  $\bar{w}$  of 153 days and the locally measured  $\epsilon_2$  of 0.0169, HIV incidence point estimates (per 100 people per year) varied between 3.03 (95% CI 2.44–3.63;

McDougal formula) and 3.19 (95% CI 2.57–3.82; Hargrove formula) (Table 2). The cBED assay-based HIV incidence estimates were thus very similar in magnitude and did not differ significantly from the estimates based on longitudinal measurement (crude and sex-age adjusted) (Table 2). Furthermore, when we implemented the cBED assay formulae using the lower bound or upper bound of the 95% CI of the locally measured long-term FPR (0.0100–0.0266), the cBED assay-based HIV incidence estimates did not differ significantly from the estimates based on longitudinal measurement. By contrast, when we implemented the cBED assay formulae using the externally measured long-term FPR of 0.0560 [6], all four cBED assay-based HIV incidence estimates were significantly lower than the longitudinal estimates (Table 2).

Our finding that the cBED assay-based HIV incidence estimate was not significantly different from the longitudinal HIV incidence

**Table 2.** HIV incidence estimates.

Estimation type	Unit	HIV incidence	
		Mean	95% CI
<b>Longitudinal measurement</b>			
(7,685 person-years, 224 seroconversions)			
Crude	(per 100 person-years)	2.91	2.56–3.32
Sex-age adjusted	(per 100 person-years)	3.14	2.73–3.58
Crude	(per 100 people per year)	2.87	2.53–3.27
Sex-age adjusted	(per 100 people per year)	3.09	2.69–3.52
<b>cBED assay measurement</b>			
(n = 11,755)			
<i>Mean of locally measured long-term FPR (<math>\epsilon_2 = 0.0169</math>)</i>			
McWalter/Welte	(per 100 person-years)	3.22	2.57–3.87
McWalter/Welte	(per 100 people per year)	3.17	2.54–3.80
McDougal	(per 100 people per year)	3.03	2.44–3.63
Hargrove	(per 100 people per year)	3.19	2.57–3.82
McDougal, simplified	(per 100 people per year)	3.12	2.51–3.73
<i>Lower bound of 95% CI of locally measured long-term FPR (<math>\epsilon_2 = 0.0100</math>)</i>			
McWalter/Welte	(100 person-years)	3.65	3.00–4.32
McWalter/Welte	(per 100 people per year)	3.58	2.95–4.22
McDougal	(per 100 people per year)	3.40	2.82–4.00
Hargrove	(per 100 people per year)	3.57	2.95–4.19
McDougal, simplified	(per 100 people per year)	3.52	2.91–4.14
<i>Upper bound of 95% CI of locally measured long-term FPR (<math>\epsilon_2 = 0.0266</math>)</i>			
McWalter/Welte	(100 person-years)	2.60	1.96–3.27
McWalter/Welte	(per 100 people per year)	2.57	1.94–3.22
McDougal	(per 100 people per year)	2.49	1.89–3.11
Hargrove	(per 100 people per year)	2.63	1.99–3.29
McDougal, simplified	(per 100 people per year)	2.53	1.92–3.17
<i>Externally measured long-term FPR (<math>\epsilon_2 = 0.0560</math>)</i>			
McWalter/Welte	(100 person-years)	0.65	0.00–1.33
McWalter/Welte	(per 100 people per year)	0.65	0.00–1.32
McDougal	(per 100 people per year)	0.66	0.00–1.33
Hargrove	(per 100 people per year)	0.71	0.00–1.43
McDougal, simplified	(per 100 people per year)	0.65	0.00–1.32

CI = confidence interval, FPR = false-positive ratio.  
doi:10.1371/journal.pone.0003640.t002

estimate did not change when we applied the window periods of 180 and 187 days (and their corresponding long-term FPRs of 0.0182 and 0.0177 (see Table 1)). Using the McWalter/Welte formula, the cBED assay-based HIV incidence was estimated at 2.63 per 100 people per year (95% CI 2.10–3.18) with a 180-day window period and at 2.56 per 100 people per year (95% CI 2.04–3.08) with a 187-day window period. Neither of these estimates was significantly different from the longitudinally measured HIV incidence estimates or from the cBED assay-based incidence estimates based on a 153-day window period (see Table 2).

As described above, we conducted sensitivity analysis of the longitudinally measured HIV incidence estimate by changing the assumption about seroconversion dates. Assuming that all seroconverters became HIV-seropositive on the day following the last negative HIV test, crude HIV incidence was estimated at 2.97 per 100 person-years (95% CI 2.61–3.39). Assuming, on the other hand, that all seroconverters became HIV-seropositive on the day of their first positive HIV test, crude HIV incidence was estimated at 2.85 per 100 person-years (95% CI 2.51–3.25). The longitudinal HIV incidence estimates were thus highly robust to changes in the approach to computing the seroconversion date. Even under the most extreme possible assumptions, the mean HIV incidence changed by only 2% of the estimate based on the mid-date assumption, as reported in Table 2.

When we stratified HIV incidence by sex and five-year age group (starting at 15 years of age), we found that none of the cBED assay-based sex and age-specific estimates differed significantly from the corresponding longitudinally measured sex and age-specific estimates. However, our samples in each of the sex-age groups were too small to detect significant differences with reasonable confidence. The coefficients of variation (CVs) of the sex-age specific cBED assay-based HIV incidence estimates ranged from 18% to 203%; in 13 of the 15 sex-age groups the CVs were larger than 25%; in 10 sex-age groups the CVs were larger than 50%; and in 4 sex-age groups they were larger than 100%.

## Discussion

In a rural community in South Africa, we found a long-term FPR of the cBED assay of 0.0169. This value is substantially lower than the two previous estimates of the ratio. The first estimate (0.0560) “was based on analysis of specimens from longer-term-infected individuals not known to have clinical AIDS, opportunistic infections, or to be on treatment” in the USA [6]. The article, in which this value was published, provides neither the sample size for the measurement nor the confidence limits around the estimate [6]. Thus we cannot test whether the estimate is significantly different from the value that we measure in rural South Africa. The second estimate (0.0520) was based on specimens from 2,749 postpartum mothers enrolled in a Vitamin-A intervention trial in Zimbabwe [7]. This second estimate was significantly higher than the value measured in our study ( $p < 0.0001$ ).

Many previous studies have used the first estimate of the long-term FPR in their estimations of HIV incidence based on cross-sectional cBED assay surveys (e.g. [9,10,12,14]). In comparing cBED-based HIV incidence estimates to HIV incidence measured longitudinally in the same population, we have demonstrated that, had we used the long-term FPR of 0.0560, we would have significantly underestimated HIV incidence in this community. By contrast, using the locally measured ratio of 0.0169, we estimated an HIV incidence that does not differ significantly from the longitudinally measured incidence.

Our findings thus confirm the previous results by McDougal et al. [6] and Hargrove et al. [7] that cBED assay-based HIV

incidence estimates are not significantly different from longitudinally measured HIV incidence, when a locally calibrated long-term FPR ratio is used to adjust for the imperfect long-term specificity of the cBED assay. At the same time, we have shown for the first time that the long-term FPR differs significantly across settings. Hence, results from studies that use a long-term FPR measured in another setting should be viewed with skepticism.

We further found that the different formulae to estimate HIV incidence based on the cBED assay results, did not produce significantly different values even though they differ in their underlying assumptions, suggesting that the choice of formula may not be very important for most practical purposes. Finally, we showed that the estimates of the long-term FPR based on data from a longitudinal HIV surveillance are very robust to changes in the definition of “long-term” (i.e. the choice of the maximum BED progression time).

Our longitudinal HIV incidence estimates in this article are slightly lower than previously published estimates from the same community [20], because the current study uses a sample that is different from the one used previously. In particular, unlike in the previous study, we excluded from the sample people who were identified as members of a household in the study area, but who did not themselves live in the area. We excluded this population group (which faces a significantly higher risk of HIV acquisition than household members who live in the study area [23]), because cross-sectional cBED assay surveys usually do not trace such non-resident household members.

HIV incidence estimates by sex and age group are important for validating the cBED assay method as an approach to measure HIV incidence [7], and are an important disaggregation for health policy and planning, e.g. in order to inform the targeting of HIV prevention interventions. Our current sample lacked the statistical power to meaningfully stratify the HIV incidence estimates. As more data becomes available from our site, we will in the future analyze HIV incidence across population subgroups.

The promise of the cBED assay for HIV surveillance, program evaluation and policy making, lies in the fact that it allows HIV incidence estimation from cross-sectional samples. Cross-sectional HIV status information, however, does not permit estimation of the long-term FPR, requiring researchers to obtain this parameter independently. It is thus important that the parameters necessary for HIV incidence estimation are calibrated using data from those settings where longitudinal follow-up is available. A meta-analysis of the long-term FPR of the cBED assay may help explain why the parameter estimates differ and allow the determination of valid regional parameter estimates.

It may further be necessary to measure the long-term FPR repeatedly over time. For instance, one of the reasons why people with non-recent HIV infections are falsely classified as recently infected by the cBED assay is viral suppression due to ART [38]. In October 2004, ART started to become available through the public health services in the community in which this study took place. However, only a very small number of patients received ART during the study period. By the end of December 2005, i.e. half a year before the end of the study period, approximately 500 patients received ART through the public ART programme in the district in which this study took place. Because the HIV surveillance covers less than half of the district population, we estimate that in December 2005 less than 250 people in the surveillance area were receiving ART out of a total resident population of approximately 65,000 [39]. Future studies will need to investigate whether our locally estimated cBED long-term FPR changes with increasing ART coverage.

An alternative to using the long-term FPR in order to adjust cBED assay-based HIV incidence estimates for the presence of

people who are falsely classified as recently HIV-infected is to use additional information on time since seroconversion to identify these individuals and correct the misclassification. Information on time since seroconversion, which can be obtained in cross-sectional surveys, could be based on biological parameters that change with time since infection (such as CD4 count, total lymphocyte count, or viral load), clinical assessment (such as screening for HIV-related diseases that indicate late-stage HIV disease [40]), and screening for ART (through a question or laboratory test).

In conclusion, our study demonstrates that without a locally measured long-term FPR HIV incidence estimates based on the cBED assay may be severely biased, but that the cBED assay performs well in HIV incidence estimation, if a locally appropriate long-term FPR is used.

## References

- McDougal JS, Pilcher CD, Parekh BS, Gershdy-Damet G, Branson BM, et al. (2005) Surveillance for HIV-1 incidence using tests for recent infection in resource-constrained countries. *AIDS* 19 Suppl 2: S25–S30.
- Gregson S, Donnelly CA, Parker CG, Anderson RM (1996) Demographic approaches to the estimation of incidence of HIV-1 infection among adults from age-specific prevalence data in stable endemic conditions. *AIDS* 10: 1689–1697.
- Williams B, Gouws E, Wilkinson D, Karim SA (2001) Estimating HIV incidence rates from age prevalence data in epidemic situations. *Stat Med* 20: 2003–2016.
- DHS (2008) All surveys by country. Available: [http://www.measuredhs.com/aboutsurveys/search/search\\_survey\\_main.cfm?SrvyTp=country](http://www.measuredhs.com/aboutsurveys/search/search_survey_main.cfm?SrvyTp=country). Accessed 10 May 2007.
- Parekh BS, Kennedy MS, Dobbs T, Pau CP, Byers R, et al. (2002) Quantitative detection of increasing HIV type 1 antibodies after seroconversion: a simple assay for detecting recent HIV infection and estimating incidence. *AIDS Res Hum Retroviruses* 18: 295–307.
- McDougal JS, Parekh BS, Peterson ML, Branson BM, Dobbs T, et al. (2006) Comparison of HIV type 1 incidence observed during longitudinal follow-up with incidence estimated by cross-sectional analysis using the BED capture enzyme immunoassay. *AIDS Res Hum Retroviruses* 22: 945–952.
- Hargrove JW, Humphrey JH, Mutasa K, Parekh BS, McDougal JS, et al. (2008) Improved HIV-1 incidence estimates using the BED capture enzyme immunoassay. *AIDS* 22: 511–518.
- Wolday D, Meles H, Hailu E, Messele T, Mengistu Y, et al. (2007) Temporal trends in the incidence of HIV infection in antenatal clinic attendees in Addis Ababa, Ethiopia, 1995–2003. *J Intern Med* 261: 132–137.
- Karita E, Price M, Hunter E, Chomba E, Allen S, et al. (2007) Investigating the utility of the HIV-1 BED capture enzyme immunoassay using cross-sectional and longitudinal seroconverter specimens from Africa. *AIDS* 21: 403–408.
- Rehle T, Shisana O, Pillay V, Zuma K, Puren A, et al. (2007) National HIV incidence measures — new insights into the South African epidemic. *S Afr Med J* 97: 194–199.
- Shisana O, Rehle T, Simbayi LC, Parker W, Zuma K, et al. (2005) South African National HIV prevalence, HIV incidence, behaviour and communication survey. Cape Town: HSRC Press.
- Mermin J, Musinguzi J, Opio A, Kirungi W, Ekwaru JP, et al. (2008) Risk factors for recent HIV infection in Uganda. *JAMA* 300: 540–549.
- Jiang Y, Wang M, Ni M, Duan S, Wang Y, et al. (2007) HIV-1 incidence estimates using IgG-capture BED-enzyme immunoassay from surveillance sites of injection drug users in three cities of China. *AIDS* 21 Suppl 8: S47–S51.
- Li SW, Zhang XY, Li XX, Wang MJ, Li DL, et al. (2008) Detection of recent HIV-1 infections among men who have sex with men in Beijing during 2005–2006. *Chin Med J (Engl)* 121: 1105–1108.
- Nesheim S, Parekh B, Sullivan K, Bulterys M, Dobbs T, et al. (2005) Temporal trends in HIV type 1 incidence among inner-city childbearing women in Atlanta: use of the IgG-capture BED-enzyme immunoassay. *AIDS Res Hum Retroviruses* 21: 537–544.
- Hall HI, Song R, Rhodes P, Prejean J, An Q, et al. (2008) Estimation of HIV incidence in the United States. *JAMA* 300: 520–529.
- UNAIDS Reference Group on Estimates Modelling and Projections (2005) Statement on the use of the BED-assay for the estimation of HIV-1 incidence for surveillance or epidemic monitoring. Available: [http://data.unaids.org/pub/EPISlides/2006/Statement\\_BED\\_Policy\\_13Dec05\\_en.pdf](http://data.unaids.org/pub/EPISlides/2006/Statement_BED_Policy_13Dec05_en.pdf). Accessed 10 October 2007.
- McWalter TA, Welte A (2008) Relating recent infection prevalence to incidence with a sub-population of non-progressors. Available: <http://arxiv.org/abs/0801.3380>. Accessed 10 June 2008.
- Welz T, Hosegood V, Jaffar S, Bätzing-Feigenbaum J, Herbst K, et al. (2007) Continued very high prevalence of HIV infection in rural KwaZulu-Natal, South Africa: a population-based longitudinal study. *AIDS* 21: 1467–1472.
- Bärnighausen T, Tanser F, Gqwedde Z, Mbizana C, Herbst K, et al. (2008) High HIV incidence in a community with high HIV prevalence in rural South Africa: findings from a prospective population-based study. *AIDS* 22: 139–144.
- Africa Centre for Health and Population Studies. Available: <http://www.africacentre.ac.za/>. Accessed 12 June 2008.
- Tanser F, Hosegood V, Bärnighausen T, Herbst K, Nyirenda M, et al. (2007) Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. *Int J Epidemiol* 37: 956–962.
- Bärnighausen T, Hosegood V, Timaeus IM, Newell ML (2007) The socioeconomic determinants of HIV incidence: evidence from a longitudinal, population-based study in rural South Africa. *AIDS* 21 Suppl 7: S29–S38.
- Shafer LA, Biraro S, Nakiyingi-Miiró J, Kamali A, Ssematimba D, et al. (2008) HIV prevalence and incidence are no longer falling in southwest Uganda: evidence from a rural population cohort 1989–2005. *AIDS* 22: 1641–1649.
- Gray RH, Li X, Kigozi G, Serwadda D, Brahmbhatt H, et al. (2005) Increased risk of incident HIV during pregnancy in Rakai, Uganda: a prospective study. *Lancet* 366: 1182–1188.
- Kapiga SH, Lyamuya EF, Lwihula GK, Hunter DJ (1998) The incidence of HIV infection among women using family planning methods in Dar es Salaam, Tanzania. *AIDS* 12: 75–84.
- Celentano DD, Bond KC, Lyles CM, Eiumtrakul S, Go VF, et al. (2000) Preventive intervention to reduce sexually transmitted infections: a field trial in the Royal Thai Army. *Arch Intern Med* 160: 535–540.
- Nelson KE, Galai N, Safaean M, Strathdee SA, Celentano DD, et al. (2002) Temporal trends in the incidence of human immunodeficiency virus infection and risk behavior among injection drug users in Baltimore, Maryland, 1988–1998. *Am J Epidemiol* 156: 641–653.
- McWalter TA, Welte A (2008) On the estimation of the proportion of true recent infections using the BED capture enzyme immunoassay. Available: <http://arxiv.org/abs/0806.1297>. Accessed 10 June 2008.
- Welte A, McWalter TA, Bärnighausen T (2008) A simplified formula for inferring HIV incidence from cross-sectional surveys using a test for recent infection. *AIDS Res Hum Retroviruses* (forthcoming).
- Buchacz K, Klausner JD, Kerndt PR, Shouse RL, Onorato I, et al. (2008) HIV incidence among men diagnosed with early syphilis in Atlanta, San Francisco, and Los Angeles, 2004 to 2005. *J Acquir Immune Defic Syndr* 47: 234–240.
- Gupta SB, Murphy G, Koenig E, Adon C, Beyrer C, et al. (2007) Comparison of methods to detect recent HIV type 1 infection in cross-sectionally collected specimens from a cohort of female sex workers in the Dominican Republic. *AIDS Res Hum Retroviruses* 23: 1475–1480.
- Hu DJ, Vanichseni S, Mock PA, Young NL, Dobbs T, et al. (2003) HIV type 1 incidence estimates by detection of recent infection from a cross-sectional sampling of injection drug users in Bangkok: use of the IgG capture BED enzyme immunoassay. *AIDS Res Hum Retroviruses* 19: 727–730.
- Loschen S, Bätzing-Feigenbaum J, Poggensee G, Cordes C, Hintsche B, et al. (2008) Comparison of the human immunodeficiency virus (HIV) type 1-specific immunoglobulin G capture enzyme-linked immunosorbent assay and the avidity index method for identification of recent HIV infections. *J Clin Microbiol* 46: 341–345.
- Sakarovich C, Rouet F, Murphy G, Minga AK, Alioum A, et al. (2007) Do tests devised to detect recent HIV-1 infection provide reliable estimates of incidence in Africa? *J Acquir Immune Defic Syndr* 45: 115–122.
- Brillinger DR (1986) The natural variability of vital rates and associated statistics. *Biometrics* 42: 693–734.

## Acknowledgments

We thank Kobus Herbst, Phumzile Dlamini, Thobeka Mngomezulu, Zanomsa Gqwede and the field staff at the Africa Centre for Health and Population Studies at the University of KwaZulu-Natal, South Africa, for their work in collecting the data used in this study and the communities in the Africa Centre demographic surveillance area for their support and participation in this study. We also thank Beverley Singh and Mahlaste Maleka for conducting the cBED immunoassay tests at the National Centre for Infectious Diseases in Johannesburg, South Africa. We thank the reviewers for helpful comments that led to improvement of the manuscript.

## Author Contributions

Conceived and designed the experiments: TB AP MLN. Performed the experiments: JV NG AP. Analyzed the data: TB AW TAM FT MLN. Contributed reagents/materials/analysis tools: CW AW TAM NM JV NG FT AP. Wrote the paper: TB. Organized data collection: CW. Contributed to data collection: NM.

37. Anderson RN, Rosenberg HM (1998) Age standardization of deaths rates: implementation of the year 2000 standard. *National Vital Statistics Reports* 47: 1–20.
38. Centre for Disease Control (2007) Using the BED HIV-1 capture EIA assay to estimate incidence using STARHS in the context of surveillance in the United States. Available: <http://www.cdc.gov/hiv/topics/surveillance/resources/factsheets/pdf/bed.pdf>. Accessed 22 December 2007.
39. Herbst K, Cooke G, Bärnighausen T, Kany Kany A, Tanser F, et al. (2008) Adult mortality and antiretroviral treatment roll-out in a rural population in KwaZulu-Natal (submitted).
40. WHO (2007) WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children. Geneva: WHO.