Dramatic increases in HIV prevalence after scale-up of antiretroviral treatment: appendix

Jaffer Zaidi[1], Erofili Grapsa[1], Frank Tanser[1], Marie-Louise Newell[1,2], Till Bärnighausen[1,3*]

[1]Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Mtubatuba, South Africa

[2] Department of Population and Global Health, Harvard School of Public Health, Boston, USA

[3] University College London Institute of Child Health, London, UK

I. Africa Centre for Health and Population Studies

Data for this analysis were obtained from the population surveillance system maintained by the Africa Centre for Health and Population Studies (www.africacentre.ac.za). The Africa Centre is a research centre funded by the Wellcome Trust and affiliated with the University of KwaZulu-Natal. Since 2000, the Africa Centre has collected demographic data on all households and since 2004 HIV status data in all adults in a 434 km2 surveillance area in Umkhanyakude district, in northern KwaZulu-Natal (22). This district is largely rural and is one of the poorest in South Africa (25).

Prior to 2007, the eligible population for HIV testing in the Africa Centre surveillance comprised of all resident men aged 15-54 years and all resident women aged 15-49 years; after 2007, the eligible population consisted of all resident men and women above 15 years of age [1].-Dried blood spots for HIV testing were obtained from all participating adults through a finger prick after written, informed consent. 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).Ethics permission for the HIV surveillance and the Africa Centre Demographic Surveillance System was obtained from the Research Ethics Committee at the College of Health Sciences, University of KwaZulu-Natal.

II. HIV surveillance participation rates

Tables S1 and S2 show the HIV surveillance participation rates by sex, age group and calendar year.

Table S1: Participation rates by age group and calendar year (women)

	Calendar year							
Age group (years)	2004	2005	2006	2007	2008	2009	2010	2011
15-19	70%	43%	50%	43%	39%	37%	46%	44%
20-24	61%	33%	44%	38%	38%	38%	44%	43%
25-29	52%	25%	34%	30%	29%	30%	42%	40%
30-34	51%	26%	30%	27%	27%	28%	41%	38%
35-39	52%	32%	33%	29%	28%	30%	40%	39%
40-44	58%	35%	37%	32%	28%	28%	39%	40%
45-49	61%	38%	39%	37%	35%	34%	44%	44%

Table S2: Participation rates by age group and calendar year (men)

Age group (years)	Calendar years							
	2004	2005	2006	2007	2008	2009	2010	2011
15-19	65%	41%	43%	36%	33%	30%	38%	37%
20-24	59%	28%	34%	27%	26%	26%	30%	31%
25-29	48%	20%	31%	24%	25%	22%	24%	26%
30-34	51%	18%	26%	22%	22%	22%	23%	24%
35-39	48%	21%	26%	20%	19%	20%	25%	25%
40-44	51%	21%	28%	24%	22%	20%	26%	25%
45-49	51%	25%	26%	20%	20%	23%	27%	30%

III. HIV prevalence estimates by age group and calendar year

Tables S3 and S4 show the HIV prevalence estimates by sex, age group and calendar year after multiple imputation to control for selection effects.

Table S3: HIV prevalence estimates in women by five-year age group (imputation-adjusted)

Age group (years)	Calendar years							
	2004	2005	2006	2007	2008	2009	2010	2011
15-19	12.6%	11.1%	11.5%	12.8%	14.2%	15.3%	14.8%	14.7%
20-24	25.6%	21.1%	23.2%	23.7%	23.1%	24.4%	26.0%	26.5%
25-29	38.4%	30.9%	31.3%	32.3%	33.4%	33.6%	40.6%	38.3%
30-34	40.9%	31.9%	33.1%	36.4%	39.5%	40.2%	47.8%	47.1%
35-39	35.0%	33.0%	34.1%	37.8%	40.4%	42.4%	50.6%	50.4%
40-44	30.0%	32.0%	32.8%	39.6%	43.2%	46.6%	51.2%	49.1%
45-49	29.2%	32.7%	31.3%	37.2%	43.2%	44.9%	49.4%	50.3%

Table S4: HIV prevalence in men by five-year age group (imputation-adjusted)

Age group (years)	Calendar years							
	2004	2005	2006	2007	2008	2009	2010	2011
15-19	3.8%	5.2%	4.7%	6.0%	6.2%	7.8%	6.7%	7.0%
20-24	8.3%	9.3%	8.2%	10.2%	10.3%	11.1%	11.1%	10.2%
25-29	19.6%	15.8%	17.2%	17.6%	18.0%	18.3%	20.0%	16.0%
30-34	30.1%	22.8%	21.0%	23.0%	21.8%	27.5%	27.4%	27.3%
35-39	27.1%	23.9%	24.4%	26.5%	29.8%	32.0%	32.7%	32.0%
40-44	27.7%	24.1%	24.8%	28.9%	29.2%	35.5%	37.0%	35.8%
45-49	22.6%	25.0%	24.0%	30.8%	30.5%	40.0%	40.2%	39.1%

III. HIV prevalence estimates based on HIV status information from wider time windows

To test the robustness of our findings to increasing the information content for the calculation of the prevalence estimates for each calendar year (while decreasing the precision of the information), we have added analyses to the paper using methods similar to those introduced by Floyd et al. [2]. We used HIV status data in windows of (i) +/- one year, (ii) +/- two years, and (iii) +/- three years around the date when an individual was contacted in the HIV surveillance but refused to participate. For these three analyses, we selected HIV status data according to the following decision rules: We used the HIV status data that was closest in time to the contact date within the time respective time window. If two HIV status data points had been available in equal distance in time from the contact date, we would have used the earlier date; however, this situation never occurred. The results of these analyses demonstrate that our findings are highly robust to the addition of HIV status data available at one point in time from the same individual who at another point in time refused an HIV test. All annual HIV prevalence estimates remained essentially unchanged in all three additional analyses. **Figure S1** shows the complete-cases analyses, using information in the different time windows around the contact date; the multiple-imputation analyses using the new datasets generated results that were also essentially the same as those obtained in the baseline analyses (**Figure S2**).

Figure S1: HIV prevalence estimates using HIV status information from wider time windows (complete-case analysis)

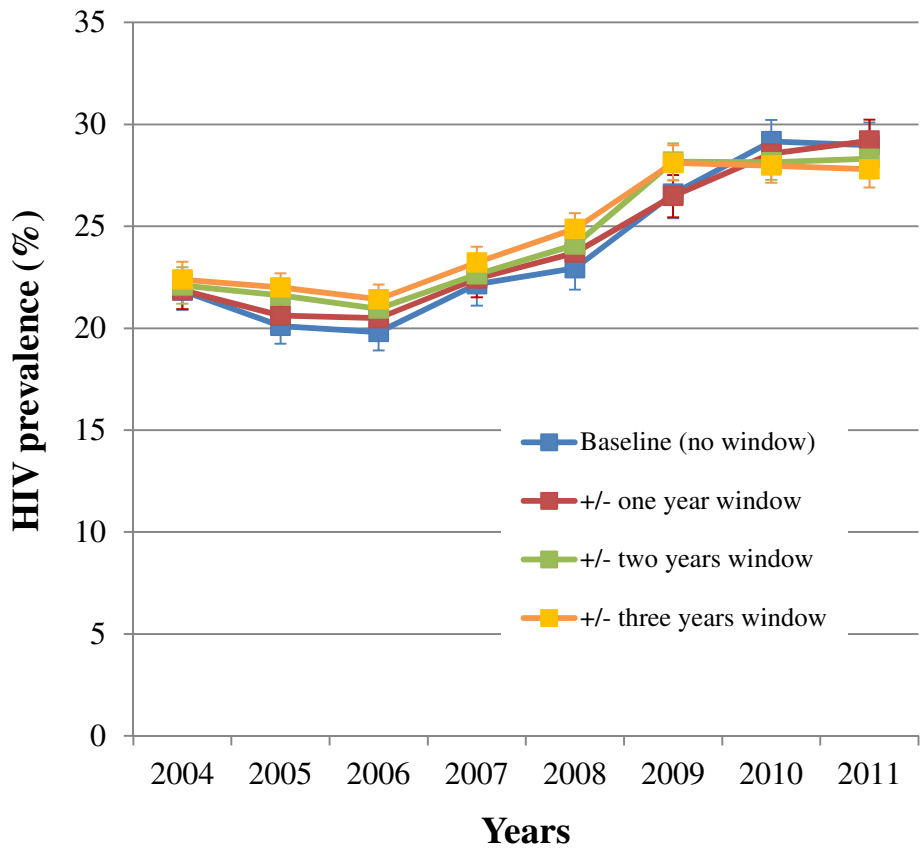
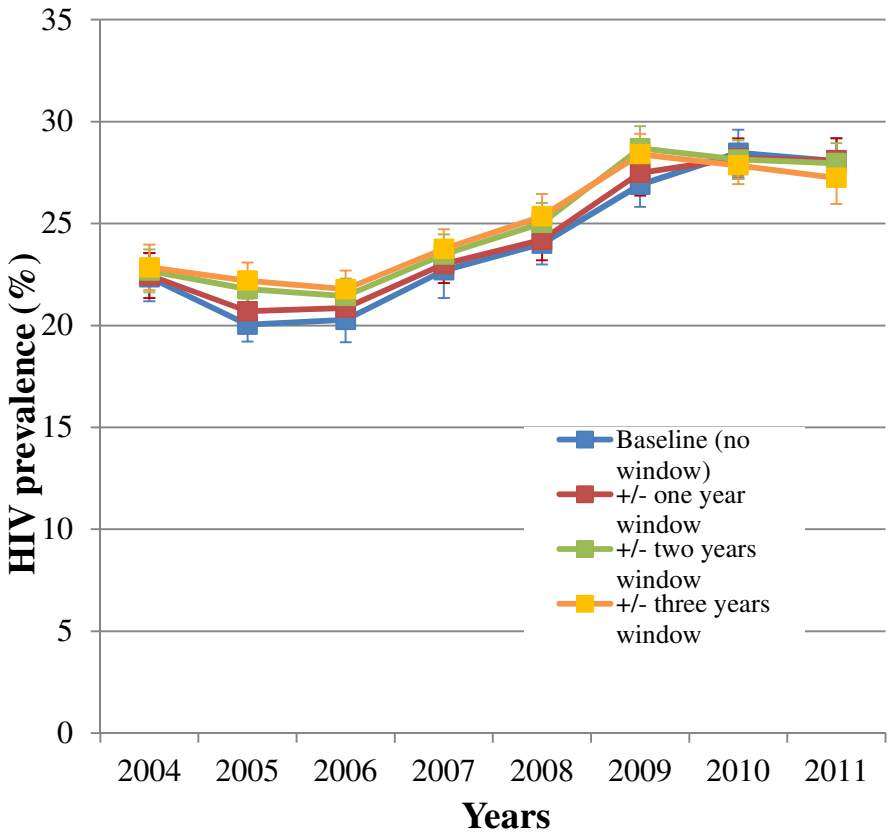


Figure S2: HIV prevalence estimates using HIV status information from wider time windows (after multiple imputation)



IV. ART coverage of all HIV-infected people

Table S5 shows estimates of HIV prevalence and ART coverage of all HIV-infected people who were eligible for HIV testing by calendar year; **Table S6** shows estimates of ART coverage of all HIV-infected people who were eligible for HIV testing by calendar year and sex and age group. The HIV prevalence trends by sex and age group vary systematically with ART coverage; both the HIV prevalence and the ART coverage increases are most rapid in women aged 25-49.

Table S5: HIV prevalence and ART coverage trends, 2004-2011

Year	Complete-case HIV prevalence estimates (95 % CI)	Imputation-adjusted HIV prevalence estimates (95% CI)	ART coverage estimates (95% CI)
2004	21.8 (20.9-22.7)	22.4 (21.2-23.5)	0.0 (0.0-0.2)
2005	20.1 (19.2-21.0)	20.0 (19.2-20.9)	1.0 (0.6-1.4)
2006	19.8 (18.9-20.7)	20.3 (19.2-21.4)	3.8 (3.2-4.6)
2007	22.2 (21.1-23.1)	22.7 (21.3-24.1)	8.3 (7.4-9.3)
2008	22.9 (21.9-24.0)	24.0 (23.0-25.0)	14.3 (13.2-15.5)
2009	26.6 (25.4-27.9)	26.9 (25.8-28.0)	20.1 (18.8-21.3)
2010	29.2 (28.1-30.2)	28.5 (27.3-29.6)	24.7 (23.4-25.9)
2011	29.0 (27.9-30.1)	28.1 (26.9-29.2)	30.7 (29.3-32.1)

ART = antiretroviral treatment, CI = confidence interval

Table S6: ART coverage of all HIV-infected people by calendar year, sex and age group

Sex	Women		Men	
Age group (years)	15-24	25-49	15-24	25-49
Calendar year	ART coverage % (95% CI)	ART coverage % (95% CI)	ART coverage % (95% CI)	ART coverage % (95% CI)
2004	0.0 (0.0-0.9)	0.3 (0.1-0.9)	0.0 (0.0-5.3)	0.0 (0-0.9)
2005	0.2 (0.0-1.3)	2.7 (1.8-4.0)	0.0 (0.0-4.6)	2.2 (1.0-4.1)
2006	1.3 (0.4-3.0)	7.8 (6.1-9.8)	0.0 (0.0-6.5)	3.6 (2.0-6.1)
2007	1.6 (0.6-3.5)	14.2 (12.0-16.8)	6.1 (1.3-16.9)	11.0 (7.8-15.1)
2008	6.1 (3.8-9.2)	21.6 (19.0-24.4)	3.8 (0.5-13.0)	16.6 (12.7-21.2)
2009	7.7 (5.0-11.1)	28.4 (25.5-31.4)	0.0 (0.0-8.4)	21.8 (17.4-26.8)
2010	10.1 (7.3-13.6)	34.6 (32.1-37.1)	5.2 (1.4-12.8)	27.5 (23.0-32.3)
2011	15.2 (11.6-19.5)	39.7 (37.0-42.5)	8.5 (2.8-18.7)	30.5 (25.0-35.7)

ART = antiretroviral treatment, CI = confidence interval

V. Multiple imputation procedure

"In order to account for selection effects on HIV participation, we used multiple imputation with chained equations. We used the R package MICE [3] to fit a logistic model on HIV status with age, sex, urban vs. rural residency status, household wealth, employment status, and educational attainment as covariates in the imputation model. The household wealth variable captures relative wealth in quintiles based on an assets index. As assets index we used the first principal component obtained in a principal component analysis of information on house ownership, water source, energy, toilet type, electricity and 27 household assets. The assets included items that can be used for consumption, production or both, such as beds, bicycles, tables, telephones, television sets, sewing machines, block makers, wheelbarrows, tractors, cattle, and other livestock. In the imputation, age and educational attainment were

used as continuous variables; urban vs. rural residency, sex, and employment status were used as categorical variables. The imputation model is given as

$$P(Y_i \mid X_i, \theta) = \text{logit}^{-1}(X_i\theta)^{Y_i} * (1 - \text{logit}^{-1}(X_i\theta))^{1-Y_i}$$

where logit^{-1}denotes the inverse of the logistic function. Maximum likelihood is used to calculate $\hat{\theta}$; the posterior variance of θ is also calculated with details given in [4]. Once, these quantities have been estimated from the fitted logistic regression, the imputation method proceeds in the following steps:

(1) Draw a θ^* from $N\left(\hat{\theta}, V\left(\hat{\theta}\right)\right)$.

(2) For each missing observation, calculate the score $\text{logit}^{-1}(X_i\theta^*)$.

(3) Compare the score to a draw from a uniform(0,1) distribution, and if the score is larger impute 1; otherwise impute 0.

Five complete datasets are generated following the procedure outlined above. From these datasets HIV prevalence and confidence intervals are generated through pooled estimation.

References

1. Tanser F, Hosegood V, Bärnighausen T, Herbst K, Nyirenda M, Muhwava W*, et al.* Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. *International Journal of Epidemiology* 2008,**37**:956-962.

2. Floyd S, Molesworth A, Dube A, Crampin AC, Houben R, Chihana M*, et al.* Underestimation of HIV prevalence in surveys when some people already know their status, and ways to reduce the bias. *AIDS* 2013,**27**:233-242.

3. Van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software* 2010:1-68.

4. Rubin DB. *Multiple imputation for nonresponse in surveys*: Wiley Online Library; 1987.