



Published in final edited form as:

AIDS. 2017 April ; 31(Suppl 1): S87–S94. doi:10.1097/QAD.0000000000001428.

Statistical Models for Incorporating Data from Routine HIV Testing of Pregnant Women at Antenatal Clinics into HIV/AIDS Epidemic Estimates

Ben Sheng¹, Kimberly Marsh², Aleksandra B. Slavkovic¹, Simon Gregson³, Jeffrey W. Eaton³, and Le Bao¹

¹Department of Statistics, Pennsylvania State University, University Park, PA, U.S.A ²Strategic Information and Monitoring Division, UNAIDS, Geneva, Switzerland ³Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom

Abstract

Objective—HIV prevalence data collected from routine HIV testing of pregnant women at antenatal clinics (ANC-RT) are potentially available from all facilities that offer testing services to pregnant women, and can be used to improve estimates of national and sub-national HIV prevalence trends. We develop methods to incorporate this new data source into the UNAIDS Estimation and Projection Package (EPP) in Spectrum 2017.

Methods—We develop a new statistical model for incorporating ANC-RT HIV prevalence data, aggregated either to the health facility level ('site-level') or regionally ('census-level'), to estimate HIV prevalence alongside existing sources of HIV prevalence data from ANC unlinked anonymous testing (ANC-UAT) and household-based national population surveys. Synthetic data are generated to understand how the availability of ANC-RT data affects the accuracy of various parameter estimates.

Results—We estimate HIV prevalence and additional parameters using both ANC-RT and other existing data. Fitting HIV prevalence using synthetic data generally gives precise estimates of the underlying trend and other parameters. More years of ANC-RT data should improve prevalence estimates. More ANC-RT sites and continuation with existing ANC-UAT sites may improve the estimate of calibration between ANC-UAT and ANC-RT sites.

Conclusion—We have proposed methods to incorporate ANC-RT data into Spectrum to obtain more precise estimates of prevalence and other measures of the epidemic. Many assumptions about the accuracy, consistency, and representativeness of ANC-RT prevalence underlie the use of these data for monitoring HIV epidemic trends, and should be tested as more data become available from national ANC-RT programs.

Keywords

Statistical Models; HIV Prevalence; Antenatal Clinic Surveillance; Routine HIV Testing Among Pregnant Women

1. Introduction

For the past two decades, HIV sentinel surveillance has been conducted based on unlinked anonymous testing among pregnant women a selection of antenatal clinics (ANC-UAT). ANCs were selected as sentinel sites for convenience and geographic spread [1]. Since the early 2000s, ANC-UAT data have often been complemented in settings with the highest burden of HIV by prevalence data from nationally-representative population-based surveys (NPS) [2]. Both data sources are important inputs for HIV prevalence models in the Estimation and Projection Package (EPP) of the Joint United Nations Programme on AIDS (UNAIDS) supported Spectrum software tool [3].

In August 2015, the World Health Organization (WHO) and UNAIDS released new guidelines recommending that countries transition from conducting ANC-UAT to using ANC routine testing (ANC-RT) data [4,5]. In the ANC-RT approach, HIV prevalence among pregnant women attending ANCs is drawn data routinely reported on a monthly or quarterly basis through existing HIV program monitoring systems. These data typically include the number of women attending their first antenatal visit, the number who are already known to be HIV positive, the number who are tested for HIV, and the number of those who test HIV positive. ANC-RT data have the potential to improve the representativeness of HIV prevalence estimates over ANC-UAT if coverage of HIV testing services at ANC and the quality and completeness of available data are consistently high. In this setting, prevalence trends are representative of all pregnant women, rather than among those from a non-representative sample of health facilities.

In anticipation of the 2015 recommendations, WHO and UNAIDS published guidelines in 2013 on how to assess and increase the use of ANC-RT data for surveillance purposes [6]. The primary focus of the assessment was to determine the quality of facility-based HIV testing services at ANCs and to provide additional support to improve services where required. The guidelines also recommended reviewing and improving processes for recording and reporting a minimum set of demographic and testing information required to interpret the estimates. Finally, countries were encouraged to explore how the switch from ANC-UAT to ANC-RT data might influence or shift HIV prevalence estimates after the transition period by comparing facility- and national-level ANC-UAT and ANC-RT data. With this final recommendation, it was recognized that in addition to the usual caveat of using ANC-UAT data to monitor general population trends, specifically, that HIV prevalence trends among pregnant women might reflect changes in the fertility patterns of infected women rather than true population trends [7, 8, 9, 10], use of ANC-RT data poses some further considerations:

1. The ‘routine’ nature of the data could lead to more errors in recording or reporting, delays or incomplete reporting by some facilities, and variation in the coverage or quality of data ascertainment over time or location.
2. HIV status ascertainment among pregnant women does not necessarily entail conducting an HIV test for every woman, but is a combination of routine HIV testing outcomes and some women self-reporting their HIV status (and ideally providing documentation).
3. Women can opt out of routine testing, which may create participation bias.
4. Inclusion of all pregnant women attending ANC results in large numbers of women tested and statistically precise estimates of prevalence, while non-sampling error arising from other sources of variability (described above) are likely to be much more important sources of uncertainty.
5. Many health facilities provide ANC-RT data, and statistically modelling individual time-series from each facility (as with existing ANC-UAT [11]) is computationally intractable.

In practice, countries have made considerable progress since 2013 in improving the coverage of testing and the availability and quality of ANC-RT data. At the same time, they have encountered challenges in quantifying some of the other possible biases described above; especially that there is little information about how the inclusion of ANC-RT estimates of HIV prevalence might influence our ability to consistently monitor trends in HIV prevalence among pregnant women. Responding to these opportunities and needs, the main objective of this paper is to describe the conceptual considerations and underlying methods used to incorporate ANC-RT data into EPP. As a secondary objective, synthetic ANC-RT data are used to illustrate how the availability of ANC-RT data affects the accuracy of various parameter estimates. In the synthetic data analysis, the availability of ANC-RT data varies in four dimensions that are 1) census-level data vs site-level data; 2) the number of sites; 3) the overlap with ANC-UAT sites; 4) the number of years of data available.

2. Method

Within Spectrum, the key element for estimating general population HIV prevalence and incidence trends is the Estimation and Projection Package (EPP). This package is a set of mathematical equations based on a susceptible-infected epidemiological model with CD4 progression patterns and ART utilization used to inform survival [12,13,14]. Model outputs of HIV prevalence are derived for people aged 15 to 49. Previous versions of the EPP software jointly fit the model to ANC-UAT and NPS data in a Bayesian statistical framework [11].

We propose two approaches to incorporating ANC-RT data. First, a ‘site-level’ approach in which ANC-RT data from the same clinics that historically participated in ANC-UAT are used as an extension of the prevalence time-series for each clinic, and possibly expanding the number of sites to a larger and more representative selection of clinics. Second, a ‘census’ approach in which ANC-RT data are aggregated across all health facilities where

HIV testing is offered to pregnant women within the geographic region being represented by the EPP model fit; in this approach, a single time series of HIV prevalence among pregnant women is calculated and used in the model.

2.1. Existing Statistical Models for ANC-UAT and NPS Data

The existing EPP model estimation utilizes separate likelihood functions to relate model predictions to HIV prevalence observations from NPS and ANC-UAT. Technical details of these likelihood functions are described elsewhere [11, 14]. Briefly, we assume that NPS provide an unbiased estimate of the true general population HIV prevalence ρ_b with appropriately estimated standard errors accounting for the complex survey design [14]. For ANC-UAT prevalence, we model the probit-transformed prevalence trend within each sentinel site, with site-level random effects ($b_s \sim \mathcal{N}(0, \tau^2)$) to capture differences in epidemic levels across sites [11]. A bias term α_{UAT} is incorporated to allow for a systematic bias (on the probit scale) between pregnant women HIV prevalence and the mean prevalence at ANC-UAT sites; this bias term captures the potential non-representative selection of ANC-UAT site locations. The model-prediction for ANC-UAT prevalence is adjusted to account for the expected change in the ratio ϕ_t of HIV prevalence among pregnant women to the general population as the epidemic matures [14]. Elsewhere in this supplement, Eaton and Bao describe another innovation to this likelihood to quantify and account for additional non-sampling error observed in ANC-UAT prevalence trends [15]; this approach is also useful for modelling ANC-RT prevalence.

2.2. Data from Routine Testing of Pregnant Women

For the purposes of this analysis, we assumed that data were available on an annual basis from each health facility for (1) the number of women attending their first antenatal visit, (2) the number already known to be HIV positive, (3) the number tested for HIV, and (4) the number testing HIV positive. Prevalence among pregnant women from routine data is calculated as:

$$\frac{[(2) \text{ known HIV+}] + [(4) \text{ tested HIV+}]}{[(2) \text{ known HIV+}] + [(3) \text{ total HIV tested}]}$$

We note some key assumptions underlying this calculation. First, the prevalence calculation includes self-reported status, and hence relies on the accuracy of this self-reporting. However, these women cannot be excluded from the calculation because this will systematically bias the resulting prevalence estimate. Second, some women might not have their HIV status ascertained (the difference between (1) number attending ANC first visit and the denominator), and this calculation assumes that prevalence is similar regardless whether women had their status ascertained. In well-functioning systems in which ANC-RT would be reliable, this number should be low, but may be an additional source of uncertainty.

In some settings, women may also be recorded as ‘previously known HIV-negative’ based on presenting documentation of a recent HIV test; in this case, the prevalence calculation is revised accordingly. Reporting systems also typically capture HIV testing of women at labor and delivery and in postpartum care. Capturing these diagnoses is important for

programmatic monitoring, but for surveillance of epidemic trends, we recommend restricting the prevalence calculation to HIV status ascertainment at the first ANC visit for continuity of the previous ANC-UAT sentinel population, because facility attendance for labor and delivery is often lower than ANC attendance and potentially a more selected and biased population, and because women may attend different facilities for delivery than for ANC, increasing the potential for selective double counting of some women.

2.3. Statistical Model for Site-Level ANC-RT Data

In countries that are currently expanding or improving ANC-RT services available to pregnant women, it may be advantageous to focus on collecting data from a limited number of health facilities where high quality data can be assured, as opposed to taking a census approach where data quality could be more variable and prevalence trends may be reflective of changes in the composition of testing sites as the program expands.

The models of site-level ANC-RT data include many of the same terms from the model of ANC-UAT data. In addition, we introduce the calibration parameter β_{RT} between ANC-UAT prevalence and ANC-RT prevalence from the same sites; this allows for the potential of a systematic bias between ANC-UAT and ANC-RT due to testing procedure differences and non-consent. The error term of site-level ANC-RT ε_{st}^{SRT} includes the sampling error with variance σ_{st}^2 and the non-sampling error. We assume that the non-sampling errors of the site-level ANC-RT and the site-level ANC-UAT have the same variance δ_{SP}^2 because both use pre-selected sites and have approaches that incorporate extensive quality assurance checks. This leads to the following site level ANC-RT data model:

$$\begin{aligned} \text{Site RT: } W_{st} &= \Phi^{-1}(\rho_t * \varphi_t) + \alpha_{UAT} + \beta_{RT} + b_s + \varepsilon_{st}^{SRT}, \\ b_s &\sim N(0, \tau^2), \\ \varepsilon_{st}^{SRT} &\sim N(0, \sigma_{st}^2 + \delta_{SP}^2), \end{aligned}$$

where the free parameters are ρ_t , α_{UAT} , β_{RT} , τ^2 , and δ_{SP}^2 .

2.4. Statistical Model for Census-Level ANC-RT Data

Census-level ANC-RT data are an aggregation of routine data from all facilities that offer ANC-RT within the geographic region represented by an EPP model fit. If ANC-RT data have sufficiently high quality and HIV testing coverage is consistently high across facilities, then this approach overcomes the non-representativeness of ANC-UAT sites. Considering the routine nature of the data, some sites may have lower quality data due to increased errors in testing or recording; however, with the large number of sites, the impact on aggregate prevalence may be small. Let α_{RT} be the calibration parameter for ANC-RT data, and ε_t^{CRT} be the error term that includes sampling error σ_t^2 and non-sampling error from census-level routine testing δ_C^2 . The model of census-level ANC-RT data includes some terms from the ANC-UAT model, but these data are no longer at site-level:

$$\begin{aligned} \text{Census RT: } W_t &= \Phi^{-1}(\rho_t * \varphi_t) + \alpha_{RT} + \varepsilon_t^{CRT}, \\ \varepsilon_t^{CRT} &\sim N(0, \sigma_t^2 + \delta_C^2), \end{aligned}$$

where the free parameters are ρ_b , α_{RT} , and δ_C^2 .

2.5. Simulation Study

We created synthetic datasets to understand how the incorporation of ANC-RT data in EPP affects estimates of the HIV prevalence trend and other parameters. We used the EPP r-spline model to simulate a single epidemic curve representative of a high HIV prevalence setting in sub-Saharan Africa (SSA). Taking this epidemic prevalence to be the ‘true’ prevalence trend, we simulated synthetic datasets consisting of two NPS occurring in 2004 and 2010, ANC-UAT at 11 sentinel sites in 1994–1996, 1998–1999, 2001, 2003, 2005, 2007, and 2010 (6 ANC-UAT sites only have data in 2007 and 2010), and ANC-RT data at 500 health facilities starting from 2011. This pattern of NPS and ANC-UAT data availability is typical of many countries in SSA.

To generate the synthetic data, we use “true” parameter values. We have “true” site-effects b_s for the 11 ANC-UAT sites and “true” site-effect variance $\tau^2 = 0.1167$. The “true” value of the ANC-UAT calibration parameter α_{UAT} is set to be 0.2402 (this corresponds to a prevalence increase from 6.00% to 9.43%) reflecting an assumption that ANC-UAT prevalence tends to be higher than general population prevalence because of the selection bias in sentinel site locations [7]. For β_{RT} , the calibration parameter between ANC-UAT prevalence and ANC-RT prevalence from the same sites, we set the “true” value at -0.1 , reflective of an assumption of a systematic difference between routine testing prevalence and that which would be obtained through ANC-UAT at the same facility; the value $\beta_{RT} = -0.1$ corresponds to, for example, a prevalence decrease from 6.00% to 4.90%. For the site-level non-sampling variance parameter δ_{SP}^2 , the “true” value is set to 0.01301. To generate synthetic census-level ANC-RT data, we aggregate the simulated site-level ANC-RT data from 500 sites.

In our synthetic data analysis, the availability of ANC-RT data varies in four dimensions that are 1) census-level vs site-level; 2) the number of sites; 3) the overlap with ANC-UAT sites; 4) the number of data years. From the first three dimensions, we have five different scenarios for the availability of ANC-RT data based on realistic cases of ANC-RT usage. These scenarios include no ANC-RT data, ANC-RT sites solely as a continuation of existing ANC-UAT sites, additional ANC-RT sites with continuation of existing ANC-UAT sites, entirely new ANC-RT sites chosen to be more geographically representative, and census-level ANC-RT data.

1. “No ANC-RT” – No ANC-RT (ANC-UAT and NPS only)
2. “11 Original” – 11 ANC-RT sites as a continuation of ANC-UAT sites,
3. “50 with Continuity” – 50 ANC-RT sites with 11 sites as a continuation of ANC-UAT sites,

4. “50 Resampled” – 50 ANC-RT sites with no continuation of ANC-UAT sites,
5. “Census” – 500 ANC-RT sites aggregated to census-level.

For each scenario except “No ANC-RT”, we vary the number of ANC-RT data years at 1, 3, and 5 years. Further details of the simulation procedure are provided in Appendix A1.

After simulating ANC-RT, ANC-UAT, and NPS data, we fit the EPP r-spline and r-trend models to the synthetic datasets, and estimate the HIV prevalence and additional parameters. During estimation, we used prior distributions including $\alpha_{UAT} \sim \mathcal{N}(0.15, 1.0)$ [14], $\beta_{RT} \sim \mathcal{N}(0, 1.0)$, $\alpha_{RT} \sim \mathcal{N}(0, 1.0)$,

$\delta_{SP}^2 \sim Exponential(mean=0.015)$, $\delta_C^2 \sim Exponential(mean=0.015)$. To evaluate the prediction accuracy, we calculate the mean absolute error (MAE) for population HIV prevalence in 2011 and 2016 (ρ_{2011} , ρ_{2016}), the year-on-year change in prevalence between 2010 to 2011 and 2015 to 2016 ($\rho_{2011} - \rho_{2010}$, $\rho_{2016} - \rho_{2015}$), and the additional model parameters α_{UAT} , β_{RT} , and δ_{SP}^2 . MAE is defined as the absolute difference averaged over 50 simulations between the median fitted value (of 3000 posterior samples) and the “true” value.

3. Results

Under the various settings of the EPP model, ANC-RT scenario, and number of ANC-RT data years, we compare the estimated and “true” values of certain quantities. We present only the results for the r-spline model [16] in this section, but many findings remain the same when we analyze the synthetic data by using the r-trend model [17] (see r-trend results in Appendix A2).

For adults, the “true” values of HIV prevalence in 2011 (ρ_{2011}) and 2016 (ρ_{2016}) were 6.62% and 5.74%, respectively. For HIV prevalence change, the “true” value of the change from 2010 to 2011 was -0.193% , and the “true” value of the change from 2015 to 2016 was -0.179% .

3.1. Site-Level ANC-RT with Continuation from ANC-UAT Sites

Table 1 presents results for “No ANC-RT” and the two site-level ANC-RT scenarios (11 and 50 ANC-RT sites) with continuation from ANC-UAT sites. With more years of ANC-RT data, for prevalence, prevalence change, and δ_{SP}^2 , the MAE generally decreases, but the mean difference did not always improve. For example, for estimation of ρ_{2016} with 11 ANC-RT sites, the MAE decreases from 0.378% for “No ANC-RT” and 0.378% for 1 ANC-RT data year to 0.364% for 5 ANC-RT data years; however, the mean difference (bias) increases from -0.052% for “No ANC-RT” and -0.079% for 1 ANC-RT data year to 0.116% for 5 ANC-RT data years. With more years, the MAE may decrease for β_{RT} , and increase for α_{UAT} .

Comparing the two site-level ANC-RT scenarios, the scenario with 50 ANC-RT sites has lower MAE and mean difference for β_{RT} . Increasing from 11 to 50 ANC-RT sites, the MAE decreases from 0.0364 to 0.0331 for 1 ANC-RT data year, from 0.0286 to 0.0262 for 3

ANC-RT data years, and from 0.0309 to 0.0276 for 5 ANC-RT data years; the mean difference decreases in magnitude from -0.0205 to -0.0098 for 1 ANC-RT data year, from -0.0229 to -0.0191 for 3 ANC-RT data years, and from -0.0242 to -0.0211 for 5 ANC-RT data years. The scenario with 50 ANC-RT sites has higher MAE for α_{UAT} , and generally lower MAE for δ_{SP}^2 (except for 1 ANC-RT data year). For prevalence and prevalence change, with more ANC-RT sites, the MAE generally does not change much.

3.2. Census-Level ANC-RT

Table 2 presents results for the “No ANC-RT” scenario and the census-level ANC-RT scenario. With more years of ANC-RT data, for HIV prevalence and prevalence change, the MAE generally decreases, but the mean difference did not always improve. For example, for estimation of ρ_{2016} , the MAE decreases from 0.378% for “No ANC-RT” and 0.382% for 1 ANC-RT data year to 0.360% for 5 ANC-RT data years; however, the mean difference increases from -0.052% for “No ANC-RT” and -0.057% for 1 ANC-RT data year to 0.113% for 5 ANC-RT data years. For α_{UAT} and δ_{SP}^2 , the MAE generally increases with more years of ANC-RT data.

3.3. Comparing Site-Level ANC-RT with and without Continuation from ANC-UAT Sites

Table 3 presents results for the two site-level ANC-RT scenarios with 50 ANC-RT sites. With more years of ANC-RT data, for HIV prevalence, prevalence change, and δ_{SP}^2 , the MAE generally decreases, but the mean difference did not always improve. For example, for estimation of ρ_{2016} without continuation from ANC-UAT sites, the MAE decreases from 0.391% for 1 ANC-RT data year to 0.365% for 5 ANC-RT data years; however, the mean difference changes from -0.093% for 1 ANC-RT data year to 0.110% for 5 ANC-RT data years. With more years of ANC-RT data, the MAE may decrease for β_{RT} , and increase for α_{UAT} .

In general, having continuation from ANC-UAT sites does not significantly change MAE for prevalence and prevalence change. For β_{RT} , continuation improves the MAE and mean difference. The MAE decreases from 0.0602 to 0.0331 for 1 ANC-RT data year, from 0.0539 to 0.0262 for 3 ANC-RT data years, and from 0.0494 to 0.0276 for 5 ANC-RT data years; the mean difference decreases in magnitude from 0.0438 to -0.0098 for 1 ANC-RT data year, from 0.0352 to -0.0191 for 3 ANC-RT data years, and from 0.0281 to -0.0211 for 5 ANC-RT data years. However, for α_{UAT} , continuation increases the MAE. For δ_{SP}^2 , having continuation from ANC-UAT sites reduces MAE for 1 ANC-RT data year, but not necessarily for 3 and 5 ANC-RT data years.

4. Discussion

In this paper, we propose statistical models that incorporate the newly recommended ANC-RT data into Spectrum software. The new set of models allow joint analysis of HIV prevalence data from NPS, ANC-UAT, and ANC-RT, so that Spectrum can better inform estimates of prevalence and other measures of the epidemic. We also applied these models to synthetic datasets, and examined how availability of ANC-RT data affects the accuracy of

various parameter estimates. Based on the synthetic data results, our conclusions are as follows.

Fitting HIV prevalence trends using synthetic data generally gives precise estimates (low mean absolute error (MAE)) of the underlying trend and other parameters. Based on our simulation results, the proposed models for ANC-RT data are appropriate for use with either the r-spline or the r-trend model in EPP.

With more years of ANC-RT data, our estimates of HIV prevalence, prevalence change, and site-level non-sampling variance became more precise as represented by the lower MAE values. (In the census-level ANC-RT scenario, the estimate of site-level variance did not improve because, in that scenario, this parameter was the variance of ANC-UAT data, not ANC-RT data.) While the mean difference sometimes improved with more ANC-RT data, in many cases, the reduction in MAE was likely due to reduced variance of the prediction when more recent ANC-RT data was incorporated. In terms of the calibration parameters, we generally obtain good estimates from only one year of ANC-RT data.

Increasing the number of ANC-RT sites (but keeping the same number of continuation sites from ANC-UAT) reduced the MAE and the mean difference for β_{RT} ; perhaps more ANC-RT sites leads to a more accurate average of ANC-RT prevalence, and hence a better estimate of the ANC-RT calibration. However, the MAE for α_{UAT} generally increases; since the mean difference for α_{UAT} generally improves, the MAE increase is due to additional uncertainty in α_{UAT} .

Having ANC-RT continuation from ANC-UAT sites gives a more precise estimate of the site-level ANC-RT calibration parameter β_{RT} ; this makes intuitive sense because β_{RT} measures the systematic difference between ANC-RT and ANC-UAT prevalence at the same sites. However, continuation also increases the MAE for α_{UAT} ; since the mean difference for α_{UAT} improves, the MAE increase is due to additional uncertainty in α_{UAT} .

When we only have a few years of ANC-RT data, the variance of non-sampling errors may not be accurately estimated. An overestimated variance parameter will make the ANC-RT data contribution negligible. Considering this, we assume the variance of ANC-RT non-sampling errors is the same as the variance of ANC-UAT non-sampling errors, since they are both collected from pregnant women. Users of EPP may choose to estimate ANC-RT and ANC-UAT variances separately, if there are sufficient years of ANC-RT data available, and Appendix A3 further investigates this.

As more high quality ANC-RT data become available, it is essential to apply the proposed model to real datasets and evaluate the many as-yet un-validated assumptions underlying our understanding of ANC-RT prevalence data. Some of these key assumptions include that: (1) trends in ANC-RT prevalence accurately reflect the HIV prevalence trends among pregnant women rather than changes in patterns of service provision or utilization, (2) self-reported HIV status by known HIV-positive women are reliable and when combined with routine testing prevalence result in accurate estimates of the true prevalence among pregnant women, (3) we are able to adequately account for changes in fertility among HIV-positive women in order to distinguish between true epidemic changes versus changes in fertility

patterns (in particular the effects of ART on fertility and fertility intentions [10]), and (4) that our proposed approaches for accounting for additional uncertainty about routine prevalence data are appropriate and can be estimated from available data. Overall, routine prevalence data provide an opportunity to substantially improve the precision and granularity of HIV epidemic estimates. However, just as programs, surveillance, and reporting systems must be continually evaluated and improved, the interpretation and modelling of these must also be reviewed and validated.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the Joint United Nations Programme on HIV/AIDS, NIH – R56 AI120812-01A1, NIH – UL1 TR000127 and TR002014, NSF IGERT – DGE-1144860, NSF – BCS-0941553. The authors are grateful to Tim Hallett, Kelsey Case, Sabrina Lamour, Jacob Dee, Peter Young, Ray Shiraishi, Peter Ghys, Tim Brown, John Stover, Andreas Jahn, and Thoko Kalua.

References

1. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Guidelines for conducting HIV sentinel serosurveys among pregnant women and other groups. Joint United Nations Programme on HIV/AIDS (UNAIDS) and World Health Organization (WHO); 2003.
2. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Monitoring HIV impact using population-based surveys. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2015.
3. Case KK, Hallett TB, Gregson S, Porter K, Ghys PD. Development and future directions for the Joint United Nations Programme on HIV/AIDS estimates. *AIDS*. 2014; 28(Suppl 4):411–414.
4. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Guidelines for conducting HIV surveillance among pregnant women attending antenatal clinics based on routine programme data. World Health Organization; 2015.
5. World Health Organization. Consolidated guidelines on HIV testing services 2015. World Health Organization; 2015.
6. UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance. Guidelines for assessing the utility of data from prevention of mother-to-child transmission (PMTCT) programmes for HIV sentinel surveillance among pregnant women. World Health Organization; 2013.
7. Gouws E, Mishra V, Fowler TB. Comparison of adult HIV prevalence from national population-based surveys and antenatal clinic surveillance in countries with generalised epidemics: implications for calibrating surveillance data. *Sexually Transmitted Infections*. 2008; 84(Suppl 1):17–23. [PubMed: 17728340]
8. Eaton JW, Rehle TM, Jooste S, Nkambule R, Kim AA, Mahy M, Hallett TB. Recent HIV prevalence trends among pregnant women and all women in sub-Saharan Africa: implications for HIV estimates. *AIDS*. 2014; 28(Suppl 4):507–514.
9. Marsh K, Mahy M, Salomon JA, Hogan DR. Assessing and adjusting for differences between HIV prevalence estimates derived from national population-based surveys and antenatal care surveillance, with applications for Spectrum 2013. *AIDS*. 2014; 28(Suppl 4):497–505.
10. Yeatman, S., Eaton, JW., Beckles, Z., Benton, L., Gregson, S., Zaba, B. Impact of ART on the fertility of HIV-positive women in sub-Saharan Africa. *Tropical Medicine and International Health*; 2016.
11. Alkema L, Raftery AE, Clark SJ. Probabilistic projections of HIV prevalence using Bayesian melding. *Annals of Applied Statistics*. 2007; 1:229–248.

12. Ghys PD, Brown T, Grassly NC, Garnett G, Stanecki KA, Stover J, Walker N. The UNAIDS estimation and projection package: A software package to estimate and project national HIV epidemics. *Sexual Transmitted Infections*. 2004; 80(Suppl 1):5–9.
13. Stover J, Brown T, Marston M. Updates to the Spectrum/Estimation and Projection Package (EPP) model to estimate HIV trends for adults and children. *Sexually Transmitted Infections*. 2012; 88(Suppl 2):11–16.
14. Brown T, Bao L, Eaton JW, Hogan DR, Mahy M, Marsh K, Mathers BM, Puckett R. Improvements in prevalence trend fitting and incidence estimation in EPP 2013. *AIDS*. 2014; 28(Suppl 4):415–426.
15. Eaton JW, Bao L. Accounting for non-sampling error in estimates of HIV epidemic trends from antenatal clinic sentinel surveillance. submitted in this supplement.
16. Hogan DR, Salomon JA. Spline-based modelling of trends in the force of HIV infection, with application to the UNAIDS Estimation and Projection Package. *Sexual Transmitted Infections*. 2012; 88(Suppl 2):52–57.
17. Bao L. A new infectious disease model for estimating and projecting HIV/AIDS epidemics. *Sexual Transmitted Infections*. 2012; 88(Suppl 2):i58–i64.

Mean difference and MAE (Mean Absolute Error) between estimated and “true” value of quantities. Results given for each data scenario at 1, 3, and 5 years of simulated ANC-RT data (except “No ANC-RT”). For “No ANC-RT”, the fit only uses simulated NPS and ANC-UAT data, and does not use simulated ANC-RT data. Each setting (of data scenario and year) uses 50 replicates under different seeds.

Table 1

Site-Level ANC-RT with Continuation from ANC-UAT Sites

Quantity	“True” Value	ANC-RT Data Scenario	Mean Difference					MAE
			Years of RT Data					
			1	3	5	1	3	
Prevalence in 2011 (ρ_{2011})	6.615%	No ANC-RT	-0.103%	-0.134%	0.103%	0.450%	0.450%	0.450%
		11 Original	-0.134%	-0.114%	0.103%	0.450%	0.467%	0.421%
		50 with Continuity	-0.167%	-0.106%	0.112%	0.480%	0.460%	0.421%
Prevalence Change 2010–2011 ($\rho_{2011}-\rho_{2010}$)	-0.193%	No ANC-RT	0.017%	0.017%	0.022%	0.022%	0.022%	0.022%
		11 Original	0.019%	0.019%	0.011%	0.022%	0.023%	0.017%
		50 with Continuity	0.020%	0.018%	0.009%	0.024%	0.022%	0.017%
Prevalence in 2016 (ρ_{2016})	5.738%	No ANC-RT	-0.052%	-0.079%	0.116%	0.378%	0.394%	0.364%
		11 Original	-0.079%	-0.067%	0.116%	0.378%	0.394%	0.364%
		50 with Continuity	-0.107%	-0.059%	0.118%	0.401%	0.388%	0.359%
Prevalence Change 2015–2016 ($\rho_{2016}-\rho_{2015}$)	-0.179%	No ANC-RT	0.008%	0.007%	0.001%	0.015%	0.014%	0.012%
		11 Original	0.008%	0.007%	0.001%	0.015%	0.014%	0.012%
		50 with Continuity	0.010%	0.007%	0.000%	0.016%	0.015%	0.013%
Calibration btw UAT and RT (β_{RT})	-0.1000	No ANC-RT	NA	NA	NA	NA	NA	NA
		11 Original	-0.0205	-0.0229	-0.0242	0.0364	0.0286	0.0309
		50 with Continuity	-0.0098	-0.0191	-0.0211	0.0331	0.0262	0.0276

Quantity	"True" Value	ANC-RT Data Scenario	Mean Difference					MAE
			Years of RT Data					
			1	3	5	1	3	
No ANC-RT								
Calibration btw UAT and pregnant women (α_{UMT})	0.2402	11 Original	-0.0171	-0.0185	-0.0350	0.0441	0.0463	0.0487
			50 with Continuity					
			0.0177	0.0115	0.0118	0.0536	0.0480	0.0597
No ANC-RT								
Non-Sampling Variance of UAT (δ_{SP}^2)	0.01301		0.00032					0.00360
11 Original								
Shared Non-Sampling Variance of UAT and RT (δ_{SP}^2)	0.01301		0.00005	0.00004	0.00033	0.00308	0.00250	0.00210
		50 with Continuity						
			0.00072	-0.00042	-0.00007	0.00322	0.00181	0.00118

Census-Level ANC-RT

Table 2

Mean difference (or mean estimate for α_{RT} and δ_C^2) and MAE (Mean Absolute Error) between estimated and “true” value of quantities. Results given for each data scenario at 1, 3, and 5 years of simulated ANC-RT data (except “No ANC-RT”). For “No ANC-RT”, the fit only uses simulated NPS and ANC-UAT data, and does not use simulated ANC-RT data. Each setting (of data scenario and year) uses 50 replicates under different seeds.

Quantity	“True” Value	ANC-RT Data Scenario	Mean Difference (or Mean Estimate)					MAE
			Years of RT Data					
			1	3	5	1	3	5
Prevalence in 2011 (p_{2011})	6.615%	No ANC-RT		-0.103%				0.450%
		Census	-0.107%	-0.120%	0.097%	0.451%	0.463%	0.419%
Prevalence Change 2010–2011 ($p_{2011}-p_{2010}$)	-0.193%	No ANC-RT		0.017%				0.022%
		Census	0.018%	0.019%	0.011%	0.022%	0.023%	0.018%
Prevalence in 2016 (p_{2016})	5.738%	No ANC-RT		-0.052%				0.378%
		Census	-0.057%	-0.072%	0.113%	0.382%	0.390%	0.360%
Prevalence Change 2015–2016 ($p_{2016}-p_{2015}$)	-0.179%	No ANC-RT		0.008%				0.014%
		Census	0.008%	0.007%	0.001%	0.015%	0.015%	0.013%
Calibration btw UAT and pregnant women (α_{UAT})	0.2402	No ANC-RT		-0.0177				0.0427
		Census	-0.0152	-0.0121	-0.0269	0.0422	0.0423	0.0436
Non-Sampling Variance of UAT (δ_{SP}^2)	0.01301	No ANC-RT		0.00032				0.00360
		Census	0.00086	0.00050	0.00078	0.00417	0.00371	0.00437
Calibration btw RT and pregnant women (α_{RT})	NA	Census	0.2315	0.2257	0.2148	NA	NA	NA
		No ANC-RT						
Non-Sampling Variance of Census RT (δ_C^2)	NA	Census	0.01455	0.00427	0.00153	NA	NA	NA
		No ANC-RT						

Table 3
Comparing Site-Level ANC-RT with and without Continuation from ANC-UAT Sites

Mean difference and MAE (Mean Absolute Error) between estimated and “true” value of quantities. Results given for each data scenario at 1, 3, and 5 years of simulated ANC-RT data. Each setting (of data scenario and year) uses 50 replicates under different seeds.

Quantity	“True” Value	ANC-RT Data Scenario	Mean Difference					MAE
			Years of RT Data					
			1	3	5	1	3	
Prevalence in 2011 (p_{2011})	6.615%	50 Resampled	-0.150%	-0.121%	0.099%	0.465%	0.442%	0.425%
		50 with Continuity	-0.167%	-0.106%	0.112%	0.480%	0.460%	0.421%
Prevalence Change 2010–2011 ($p_{2011} - p_{2010}$)	-0.193%	50 Resampled	0.019%	0.019%	0.010%	0.023%	0.023%	0.017%
		50 with Continuity	0.020%	0.018%	0.009%	0.024%	0.022%	0.017%
Prevalence in 2016 (p_{2016})	5.738%	50 Resampled	-0.093%	-0.073%	0.110%	0.391%	0.371%	0.365%
		50 with Continuity	-0.107%	-0.059%	0.118%	0.401%	0.388%	0.359%
Prevalence Change 2015–2016 ($p_{2016} - p_{2015}$)	-0.179%	50 Resampled	0.009%	0.007%	0.001%	0.015%	0.015%	0.013%
		50 with Continuity	0.010%	0.007%	0.000%	0.016%	0.015%	0.013%
Calibration btw UAT and RT (β_{RT})	-0.1000	50 Resampled	0.0438	0.0352	0.0281	0.0602	0.0539	0.0494
		50 with Continuity	-0.0098	-0.0191	-0.0211	0.0331	0.0262	0.0276
Calibration btw UAT and pregnant women (α_{UAT})	0.2402	50 Resampled	-0.0273	-0.0229	-0.0378	0.0466	0.0467	0.0533
		50 with Continuity	0.0177	0.0115	0.0118	0.0536	0.0480	0.0597
Shared Non-Sampling Variance of UAT and RT (δ_{SP}^2)	0.01301	50 Resampled	0.00125	-0.00028	-0.00008	0.00426	0.00157	0.00119
		50 with Continuity	0.00072	-0.00042	-0.00007	0.00322	0.00181	0.00118