**Estimating district-level HIV incidence among women in Malawi using antenatal care (ANC) data**

Katherine M. Jia[1,2§], Andreas Jahn[3,4], Tiwonge Chimandule[3,4], Rose Nyirenda[3], Jeffrey W Imai-Eaton[1,2]

[§] Center for Communicable Disease Dynamics, 677 Huntington Ave, Suite 506, Boston, MA 02115. kjia@g.harvard.edu. 617-495-1000

[1] Center for Communicable Disease Dynamics, Department of Epidemiology,
Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States

[2] MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK.

[3] Directorate of HIV, STI, & Viral Hepatitis, Ministry of Health, Lilongwe, Malawi

[4] International Training and Education Center for Health (I-TECH), University of Washington, Lilongwe, Malawi

## Background

Sub-Saharan Africa (SSA) has a high HIV burden with 25.6 million people living with HIV (PLHIV) and 890,000 new HIV infections in 2021, accounting for two-thirds of PLHIV and 59% of new infections worldwide.[1] Monitoring trends in HIV incidence is essential to evaluate effectiveness of HIV prevention programs and tracking progress towards the UNAIDS goal of reducing new HIV infections and AIDS-related deaths by 90% between 2010 and 2030. [2]

Population HIV incidence is challenging to estimate due to limited longitudinal and representative samples of the general public undergoing routine HIV testing. Since the early 2010s, many SSA countries have consistently routinely ascertained HIV status among all pregnant women who attend antenatal care (ANC) health facilities in order to identify women needing antiretroviral medication to prevent mother-to-child HIV transmission.[3] For example, ANC uptake rate in Malawi exceeded 98% in 2015 and HIV testing coverage during ANC visits exceeded 95% in 2017.[4,5]

Pregnant women attending ANC have long been an important sentinel population for monitoring population HIV trends because women attend ANC irrespective of HIV status or specific HIV risk factors.[6] Fitting mathematical models to HIV prevalence among pregnant women, along with national household survey data, has been the main strategy to estimate national HIV incidence trends since the 1990s.[7] The large increase in HIV data through routine HIV status ascertainment of all pregnant women provides opportunities to more granularly quantify HIV infection patterns, including subnational trends and variation (e.g. district) and with finer temporal resolution. District-level estimates are important to understand the differential impact of programmes and for resource allocation in identifying areas that have the greatest need for a tailored HIV prevention strategy and will benefit the most from it.

Using the routinely-collected HIV data from ANC facilities, we developed a mathematical model to estimate incidence in the general women aged 15 to 49 years old in 28 districts of Malawi, 2011-2020. Among SSA countries, Malawi has consistently high HIV testing coverage in the ANC facilities since 2011 (≥78%) and was selected as the first country for which we will estimate HIV incidence using our model. [8] In future, we aim to apply the model to estimate HIV incidence in other SSA countries as well.

## Methods

*Data.* We analysed routinely-collected ANC data captured by the Department of HIV and AIDS Management Information System (DHA-MIS) in Malawi. Malawi is a southeastern African country with 20.9 millions individuals living in 28 districts.[9] DHA-MIS covers all ANC health facilities with HIV services under the Ministry of Health of Malawi. [8] HIV testing data are available by quarter from 3rd quarter 2011 to 4th quarter 2020. Data consist of aggregate counts of the final HIV status for pregnant women who had their final ANC visit before birth during the quarter. HIV status was classified as HIV negative, newly diagnosed HIV positive during pregnancy, already known HIV positive before pregnancy, or HIV status unknown; among women diagnosed with HIV, it is further recorded whether they were on ART before pregnancy or initiated ART during pregnancy. We calculated the HIV prevalence among ANC attendees for each district $i$ at quarter $q$ ($ANC\ Prevalence_{i,q}$), defined by the total number of HIV positive ANC attendees ($Y_{i,q}$) divided by the total number of ANC attendees with status ascertained ($N_{i,q}$).

*Model.* We developed an age-stratified compartmental model to simultaneously track HIV acquisition and fertility processes. We fitted it to district-level ANC prevalence data from DHA-MIS to estimate yearly HIV incidence (new infections per 1,000 person-year) among women aged 15-49 in Malawi from 2011 to 2021. Model structure is shown in **Figure 1**. In brief, women transitioned from sexually non-active to active, then to pregnant and lastly post-partum compartments. At any point, they could acquire HIV and move from susceptible to infection compartments. Recently infected are those receiving a test positive test during the first six months after infection, while the undiagnosed women progress into latent infection until later diagnosis. Diagnosed patients were assumed to be on ART without dropout.

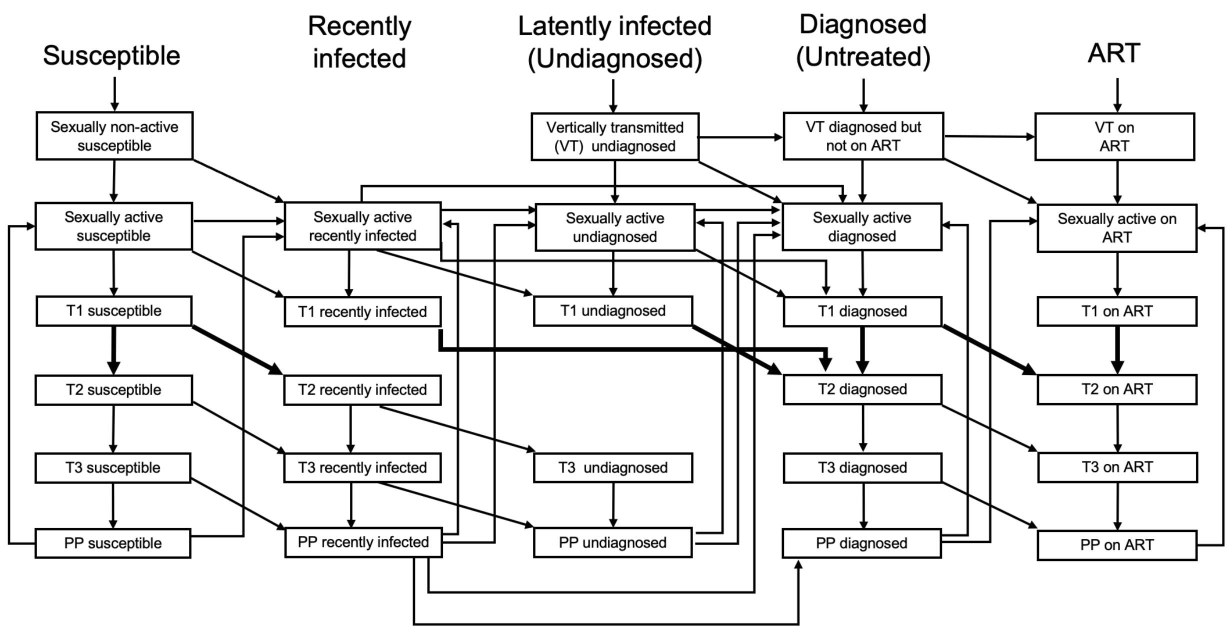


**Figure 1 | Model flowchart.** Pregnant women are assumed to attend antenatal care (ANC) facilities and test for HIV at their 2nd trimester and 2nd trimester only, as represented by the bolded arrows. We assumed that they are ANC that were observed in the DHA-MIS data.

For each district $i$ and quarter $q$, we estimated the rate of infection ($\lambda_{i,q}$) and the initial HIV prevalence in the general population ($Prev_{2011,q}$). $\lambda_{i,q}$ was assumed to follow a log linear trend with district-specific intercept $\beta_{i,0}$ and slope $\beta_{i,1}$:

$$\lambda_{i,q} = \exp\,(\beta_{i,0} + \beta_{i,1} * q\,)$$

for quarter $q$= 2011 Q3, 2011 Q4, …, 2020 Q4. We used a maximum-likelihood approach to estimate the parameters. Given an observed $N_{i,q}$ number of ANC attendees with ascertained HIV status and $Y_{i,q}$ total positives, the likelihood is

$$Y_{i,q} \sim Beta - Binomial(N_{i,q}, \alpha_{i,q}, \beta_{i,q})$$

, where $\alpha_{i,q} = \frac{\rho_{i,q}(1-\sigma_i)}{\sigma_i}$, $\beta_{i,q} = \frac{\alpha\,(1-\rho_{i,q})}{\rho_{i,q}}$, $\sigma_i$ = over-dispersion parameter, $\rho_{i,q}$ = model-predicted $ANC\ Prevalence_{i,q}$. For the initial conditions of the model at 2011, the initial HIV prevalence $Prev_{2011,q}$ were estimated from the data, and the proportion sexually debuted was obtained by fitting a skewed logistic model to the Demographic and Health Survey (DHS) 2017 indicator.[10] For the compartments within the HIV-positive, sexually active women, we assumed that the distribution of

PLHIV was in equilibrium when the model was initialized in 2011. Similarly, for the compartments within HIV-negative women, the distribution is assumed to be at equilibrium in 2011.

**Results**

Figure 2A displays the observed HIV prevalence among pregnant women at ANC by quarter in three selected districts in southern Malawi, and the HIV prevalence trend among pregnant women from modelled results. Figure 2B displays the corresponding HIV incidence estimated by the model for women aged 15-49 years, with comparison to national incidence estimates derived from the Spectrum model and reported by UNAIDS.[11] The model reflected the trend in ANC prevalence over time and the model generated prevalence is similar to the observed data in magnitude.

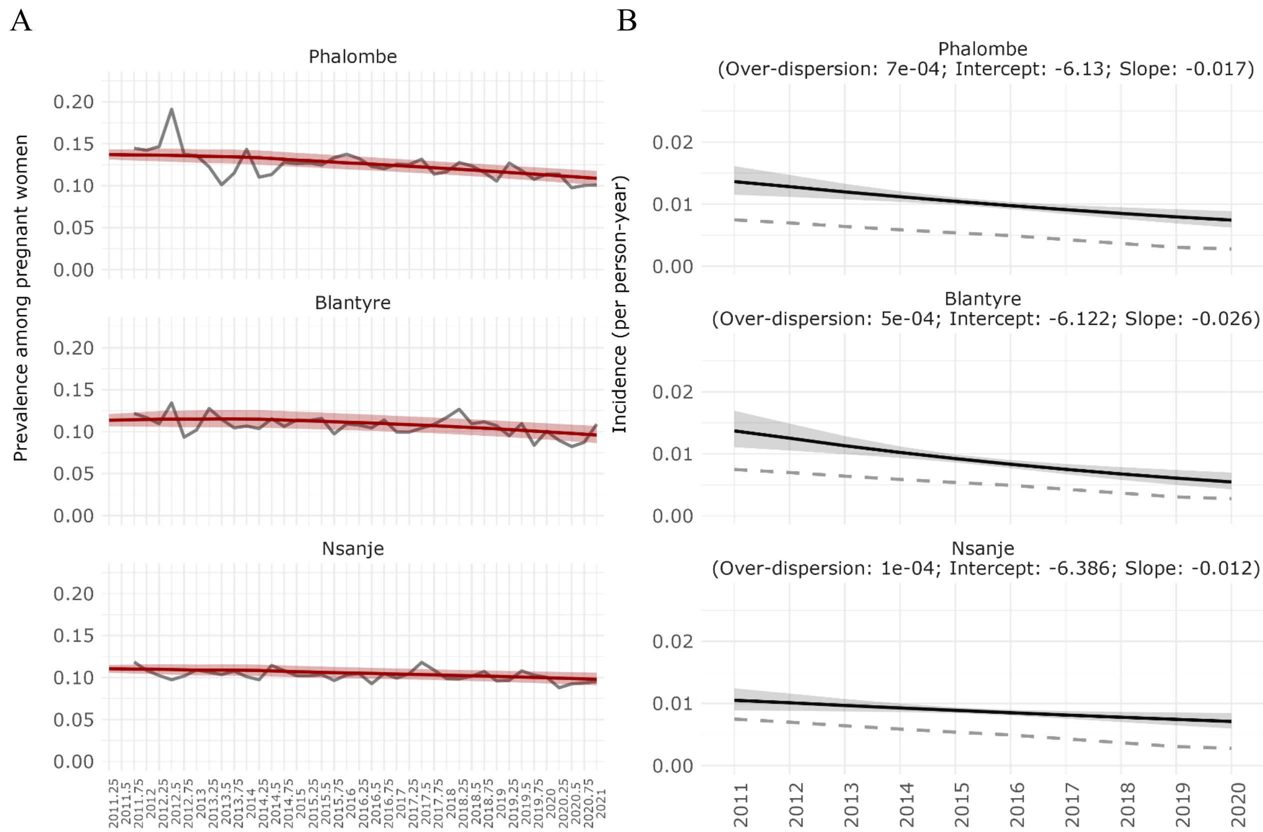


**Figure 2** | (A) Modelled HIV prevalence among pregnant women (red) compared to observed ANC prevalence (black); (B) Modelled HIV incidence (solid) compared to nationwide HIV incidence estimate from the Spectrum model (dash) [12]

Figure 3 shows the estimated incidence in all districts in 2011 and 2020. Incidence is heterogeneous across districts. For example, the district with the highest incidence (Mulanje) in 2020 has more than 20-fold incidence than the district with lowest incidence (Nkhotakota). The three regions have starkly different incidence: in 2020, the estimated incidence was highest in Southern districts (grey bars) and lowest in Central districts (orange bars).

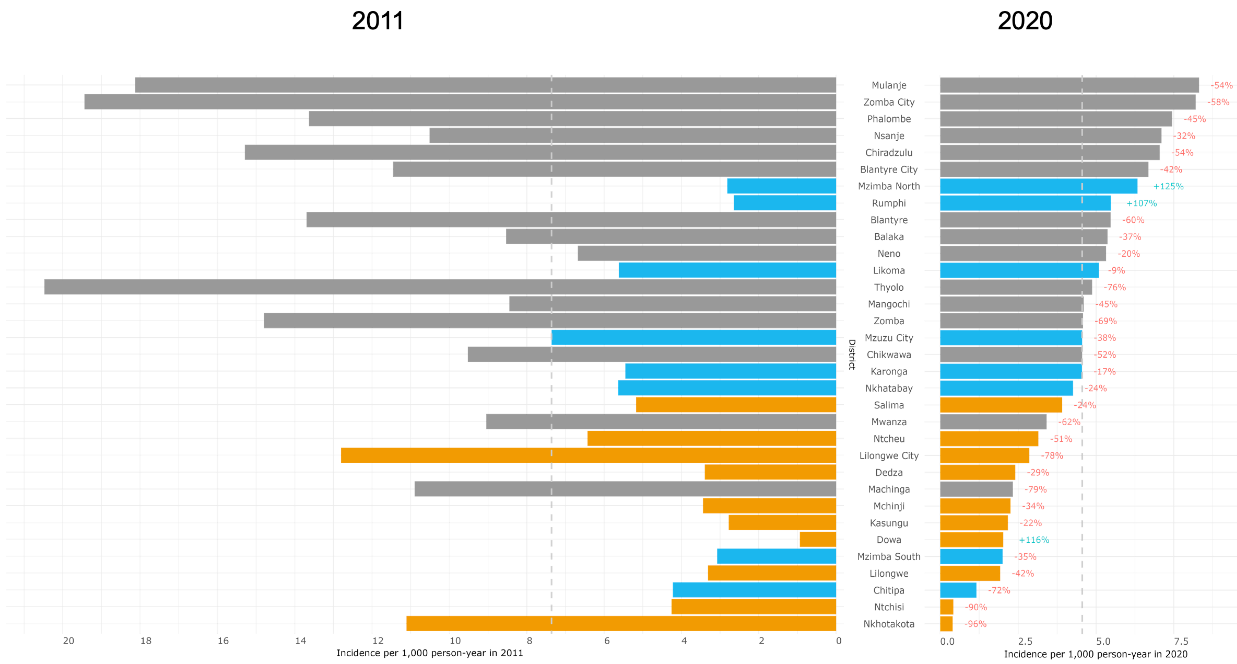


**Figure 3** | Estimated incidence among women aged 15-49 by districts in Malawi in 2011 and 2021. Districts in the northern region are colored in blue, central region in orange, and southern region in grey. Bars are labeled with the percentage change in incidence from 2011 to 2020.

Regarding incidence declines from 2011 to 2020, incidence decreased most in Southern and Central districts. Majority of Southern district decreased close to a half or more (inter-quartile range [IQR]: -40% to -60%), while the majority of Central districts decreased by a quarter or more (IQR: -25 to -70%). Majority of Northern districts also decreased but not as much as the South and Central districts, some Northern districts even increased in incidence.

Among all the districts, Nkhotakota, Ntchisi, Machinga, Liongwe City, and Thyolo, had largest declines in incidence, all of which are in the Southern or Central region. Mzimba North, Rumphi and Dowa had incidence increased.

There are several limitations of the analysis. First, the model should reflect the test-and-treat cascade under Option B+. Although this may have little impact on the current analysis using the total positives ($Y_{i,q}$) to estimate incidence, since both treated and untreated women are included in the data on $Y_{i,q}$, the lower assumed proportion on ART will impact the estimate if the model is fitted to the number of treated ANC attendees, which is one of our next steps. Second, the model was parameterized by assuming that age-specific sexual debut rate was constant for all districts. In reality, some districts may have higher sexual debut rate than others, which will impact the incidence estimate. Similarly, we do not have the number of vertically-transmitted individuals by districts and approximate the numbers of vertically-transmitted individual by scaling the national total with the relative population sizes of the districts. If a district has a high proportion of vertically-transmitted individuals (i.e., our assumption that every district has the same percentage vertically-transmitted does not hold), fitting the model to the ANC prevalence data would over-estimate the HIV incidence, since a larger percentage of the observed new positives were in fact not recently infected but were long-term survivors.

In conclusion, we estimated HIV incidence among women in Malawi using routinely-collected ANC data. Our results echo with others that incidence was higher in the Southern region. [13,14] Most districts had incidence decreased from 2011 to 2020, with reductions being more pronounced in the Southern and Central districts.

**References**


1. UNAIDS. *In Danger: UNAIDS Global AIDS Update 2022*. Accessed October 7, 2023. https://www.unaids.org/sites/default/files/media_asset/2022-global-aids-update_en.pdf

2. Stover J, Glaubius R, Teng Y, et al. Modeling the epidemiological impact of the UNAIDS 2025 targets to end AIDS as a public health threat by 2030. Rosen S, ed. *PLoS Med*. 2021;18(10):e1003831. doi:10.1371/journal.pmed.1003831

3. Dee J, Garcia Calleja JM, Marsh K, Zaidi I, Murrill C, Swaminathan M. HIV Surveillance Among Pregnant Women Attending Antenatal Clinics: Evolution and Current Direction. *JMIR Public Health Surveill*. 2017;3(4):e85. doi:10.2196/publichealth.8000

4. Maheu-Giroux M, Marsh K, Doyle C, et al. *National HIV Testing and Diagnosis Coverage in Sub-Saharan Africa: A New Modeling Tool for Estimating the "First 90" from Program and Survey Data*. Epidemiology; 2019. doi:10.1101/532010

5. National Statistical Office - NSO/Malawi and ICF. *Malawi Demographic and Health Survey 2015-16*. Accessed October 7, 2023. https://dhsprogram.com/publications/publication-fr319-dhs-final-reports.cfm

6. Johnson LF, Kubjane M, Eaton JW. Challenges in estimating HIV prevalence trends and geographical variation in HIV prevalence using antenatal data: Insights from mathematical modelling. Lima VD, ed. *PLoS ONE*. 2020;15(11):e0242595. doi:10.1371/journal.pone.0242595

7. Eaton JW, Bao L. Accounting for nonsampling error in estimates of HIV epidemic trends from antenatal clinic sentinel surveillance. *AIDS*. 2017;31(Supplement 1):S61-S68. doi:10.1097/QAD.0000000000001419

8. Maheu-Giroux M, Jahn A, Kalua T, Mganga A, Eaton JW. HIV surveillance based on routine testing data from antenatal clinics in Malawi (2011–2018): measuring and adjusting for bias from imperfect testing coverage. *AIDS*. 2019;33(Supplement 3):S295-S302. doi:10.1097/QAD.0000000000002356

9. United Nations Population Fund. *World Population Dashboard: Malawi*.; 2023. https://www.unfpa.org/data/world-population/MW

10. Nguyen VK, Eaton JW. A model for reconstructing trends and distribution in age at first sex from multiple household surveys with reporting biases. *Epidemics*. 2022;40:100593. doi:10.1016/j.epidem.2022.100593

11. UNAIDS. *AIDSinfo Global Data on HIV Epidemiology and Response*.; 2024. Accessed February 12, 2024. https://aidsinfo.unaids.org/

12. Stover J, Brown T, Puckett R, Peerapatanapokin W. Updates to the Spectrum/Estimations and Projections Package model for estimating trends and current values for key HIV indicators. *AIDS*. 2017;31(Supplement 1):S5-S11. doi:10.1097/QAD.0000000000001322

13. Wolock TM, Flaxman S, Chimpandule T, et al. *Subnational HIV Incidence Trends in Malawi: Large, Heterogeneous Declines across Space*. Epidemiology; 2023. doi:10.1101/2023.02.02.23285334

14. Eaton JW, Dwyer-Lindgren L, Gutreuter S, et al. Naomi: a new modelling tool for estimating HIV epidemic indicators at the district level in sub-Saharan Africa. *J Int AIDS Soc*. 2021;24(S5). doi:10.1002/jia2.25788