



Published in final edited form as:

J Acquir Immune Defic Syndr. 2025 April 01; 98(4): 363–371. doi:10.1097/QAI.0000000000003588.

Model-Based Prioritization of Adolescent Girls and Young Women for HIV Prevention Services Based on Data from 13 Sub-Saharan African Countries

Steve Gutreuter, PhD^a, Langan Denhard, MPH^a, Joseph E. Logan, PhD^a, Jesse Blanton, PhD^a, Haddi Jatou Cham, MPH^b

^aDivision of Global HIV and TB, U.S. Centers for Disease Control and Prevention (CDC), Atlanta, United States of America

^bDivision of Global HIV and TB, U.S. Centers for Disease Control and Prevention (CDC), Harare, Zimbabwe

Abstract

Background: Adolescent girls and young women (AGYW) aged 15–24 years are more likely to acquire HIV than their male counterparts, and well-targeted prevention interventions are needed. We developed a method to quantify the risk of HIV acquisition based on individual risk factors and population viral load (PVL) to improve targeting of prevention interventions.

Setting: This study is based on household health survey data collected in 13 sub-Saharan African countries, 2015–2019.

Methods: We developed a Bayesian spatial model which jointly estimates district-level PVL and the probability of infection among individual AGYW, aged 15–24 years, based on individual behavioral/demographic risk factors and area-level PVL. The districts (second subnational level) typically comprise the areas of estimation. The model borrows strength across countries by incorporating random effects which quantify country-level differences in HIV prevalence among AGYW.

Results: The combined survey data provided 52,171 questionnaire responses and blood tests from AGYW, and 280,323 blood samples from all respondents from which PVL was estimated. PVL was—by far—the most important predictor of test positivity (aOR = 70.6; 0.95-probability credible interval 20.7–240.5). Having a partner with HIV increased the odds of testing positive among AGYW who were never (aOR = 12.1; 7.5–19.6) and ever pregnant (aOR = 32.1; 23.7–43.4). The area under the cross-validated receiver-operating characteristic curve for classification of test positivity was 82%.

Corresponding author: Joseph E. Logan; Division of Global HIV & TB, Global Health Center, U.S. Centers for Disease Control and Prevention, Atlanta, Georgia | ffa3@cdc.gov.

Conflicts of Interest: No conflicts are declared.

Supplemental Digital Content

Supplemental_Digital_Content_1.pdf

Supplemental_Digital_Content_2.zip

Conclusion: The fitted model provides a statistically principled basis for priority enrollment in HIV prevention interventions of those AGYW most at risk of HIV infection and geographic placement of prevention services.

Keywords

infection risk; AGYW; population viral load; screening; joint modeling; errors in covariates; DREAMS; pre-exposure prophylaxis; spatial modeling

Introduction

Adolescent girls and young women (AGYW) in many sub-Saharan African (SSA) countries remain at high risk of acquiring HIV compared to their male counterparts¹. In 2022, SSA AGYW aged 15–24 years were more than three times more likely to acquire HIV than their male counterparts and accounted for over 77% of new infections among persons of this age group worldwide¹. There is substantial variation in risk and in opportunities to reach those AGYW most at risk of HIV infection^{2–4}. A UNAIDS goal is to achieve HIV epidemic control by 2030, but the current trajectory of new infections could far exceed the benchmark goal of having reduced new infections to 370,000 incidents in 2025⁵. It will be paramount for HIV prevention and treatment efforts to reach those AGYW who are most vulnerable to HIV infection in SSA if the goal of reaching epidemic HIV control is to become a reality.

The President's Emergency Plan for AIDS Relief (PEPFAR) funds programs focused on reducing HIV infection among AGYW in SSA. Such efforts include the Determined, Resilient, Empowered, AIDS-free, Mentored and Safe (DREAMS) and the Orphans and Vulnerable Children (OVC) programs^{6,7}, and also increasing uptake of pre-exposure prophylaxis (PrEP). Although all these efforts have shown promise in reducing HIV risk and infections among AGYW^{8–11}, one key ongoing challenge has been identifying and reaching those AGYW who are most vulnerable to infection with HIV.

Data on district- or county-level HIV-related epidemiological indicators are most often used to inform placement of HIV prevention and treatment strategies for AGYW. Such indicators include HIV incidence rates, numbers of AGYW at-risk, availability of HIV prevention or treatment interventions and healthcare services, and prevalence of viral suppression among persons living with HIV. The predominant metric examined for program placement is district- or county-level HIV incidence, which was highly useful for placing services at the peak of the epidemic. Now that district-level HIV incidence rates have declined to low levels in most PEPFAR-supported countries, it is becoming more challenging to use this indicator to identify hotspots of HIV transmission or to describe specific groups of vulnerable people.

To improve programmatic impact, it may be time to explore other or additional methods for placing interventions and identifying those AGYW most vulnerable to HIV⁵. It is also essential to understand the most salient vulnerabilities to ensure that interventions and services are appropriately designed to address them. Past efforts have attempted to estimate the numbers of AGYW who would be considered eligible for the DREAMS program using criteria from existing programmatic guidance, and found that the majority of AGYW in priority countries met at least one criterion¹². A 2021 review of 17 AGYW HIV prevention

screening tools in use across ten countries found that the number of questions ranged from 5 to 73, indicating a need to assess the predictability of criteria in use¹³.

The use of predictive models to identify AGYW most vulnerable to HIV has potential to complement current strategies that inform the placement and content of HIV prevention services for AGYW. Efforts to both identify areas of concentration of predefined risk groups of AGYW¹⁴ and to identify individual AGYW at elevated risk^{2,15–18} have been undertaken. Among the latter, only a single study from South Africa produced an area under the receiver-operating characteristic curve (AUC) that exceeded 75%¹⁷. That model incorporated individual-level behavioral and demographic predictors, but not measures of local HIV burden. More work is needed to explore the use of predictive modeling for HIV acquisition among AGYW based on a broader spectrum of predictors relevant to individual predisposition and the interactions they have with their environment¹⁹. Models should aim to include factors associated with HIV vulnerability at the individual- (e.g., age, education, employment, substance abuse/misuse, being sexually active, orphanhood, early pregnancy), relationship- (i.e., person-to-person), transmission- (e.g., having forced or coerced sex, unable to enforce condom use, having a partner with HIV, engaging in transactional sex, experiences of sex-based violence), community/environment- (e.g., the amount of virus exposure in one's community or environment), and societal (e.g., sex norms, policies, and laws which can reduce agency among girls and women) levels. Although data are not available to include all factors in one model, we took an initial step to improve identifying AGYW at high risk of HIV by using a more comprehensive predictive model that incorporates predictors at multiple levels of the social and ecological context.

Our objective was to develop a model which predicts probabilities of acquiring HIV among the AGYW of 13 sub-Saharan African countries based on individual- and relational-level risk factors, and a key environmental risk—the severity of the local epidemic as measured by population viral load. HIV prevention resources are limited, and we propose the use of such models to identify priority areas for placement of prevention interventions and to prioritize HIV-negative AGYW having the highest probabilities of HIV infection for enrollment.

Methods

Data sources

Population-Based HIV Impact Assessment (PHIA)²⁰ surveys were our source of data for putative risk factors and HIV test results among AGYW of ages 15–24 years, and HIV viral load assay results among all survey respondents regardless of age or sex. The PHIA surveys are multi-stage stratified household-based surveys designed to estimate HIV viral-load suppression with specified precision at the stratum level. The PHIA surveys were reviewed and approved by the CDC and the institutional review boards of each partner implementing the surveys. We combined data from 13 surveys completed between 2015–2019 (Table 1). The numbers of respondents tested for HIV ranged from 11,682–31,579 per country and totaled 280,323, including 52,171 AGYW of ages 15–24 years. Additional details, including blood-sampling procedures are provided in the survey-specific manuals²⁰. Estimated HIV prevalence among AGYW in 2023 ranged from 0.3% in Ethiopia to 7.5% in Eswatini²¹.

PEPFAR-supported prevention services are typically provided at the second subnational administrative level, herein named target subnational units (TSNUs). Eswatini, Kenya and Lesotho are exceptions, where our TSNUs were the first subnational administrative units. We located the masked centroid coordinates of the PHIA clusters within TSNUs using the 2023 UNAIDS Combined Geographic Health Boundaries file (see Data Availability). The numbers of TSNUs per country ranged from 4–178 and totaled 1,044 (Table 1; see Supplemental Digital Content 2 for the complete list of TSNUs). We also used TSNU-level estimates of the total numbers of AGYW and AGYW living with HIV from Naomi²² model output for 2023 to estimate the numbers of at-risk (HIV-negative) AGYW.

Risk covariates

The selection of behavioral and demographic predictors of HIV infection is described in detail in Supplemental Digital Content 1. Briefly, we extracted and recoded responses to PHIA survey questions to produce the categorical predictors: urbanicity (urban vs rural residence), age range (15–19, 20–24 years), partner HIV status (no current partner, partner HIV negative, partner status unknown, partner HIV positive), educational attainment (never attended school or unknown, primary only, secondary, some post-secondary), sexual history (never had sex, sexual debut at age >16, sexual debut at age ≤ 16 yr), employment during past year (no, yes), history of pregnancy (no, yes), and history of transactional sex (no, yes). The covariates were selected following review by subject matter experts supporting AGYW prevention programs based on alignment with eligibility criteria measures in DREAMS program global guidance. Global DREAMS individual eligibility criteria were determined based on review of existing literature of factors associated with HIV infection among AGYW²³. Our covariates span the individual (e.g., age, education, pregnancy etc.), relational- (sex partner characteristics) and community/environment (area of residence and population viral load; see below) levels.

Joint model of HIV risk and population viral load

We developed a Bayesian spatial model which jointly estimates the probability of HIV infection among individual AGYW and the TSNU-level population HIV viral load (PVL), defined as the TSNU-level posterior mean of $\log_{10}(\text{VL} + 1)$, where VL is the number of RNA copies per ml of blood plasma. VL was measured among all HIV-positive survey participants and therefore we defined $\text{VL} \equiv 0$ for all HIV-negative participants. Individual probabilities of HIV infection among AGYW were modeled using a logistic component, which included the individual behavioral and demographic risk covariates and PVL as predictors. PVL cannot be observed directly but was instead estimated with error and included as a classical error-prone covariate^{24,25}. The possibility of spatial correlation was accommodated using TSNU-level BYM2^{26,27} spatial smoothing structures in the estimation of both HIV positivity among the AGYW and estimation of PVL. The BYM2 model extends the more popular BYM model²⁸ by enabling scaling which facilitates hyperprior specification²⁹.

It is important to accommodate, to the extent possible, features of the survey design in Bayesian model-based estimation and prediction from survey data^{30–33}. Bayesian model-based inference from survey data predicts outcomes by including the survey sampling

probabilities or weights as predictors^{31,34,35}. Our models were fitted to pooled data from 13 probability surveys and therefore we re-scaled the sampling weights so that their sums equaled the effective sample sizes for each survey³⁶. The effective sample size for a survey was the actual sample size divided by the design effect for estimation of HIV prevalence.

We fitted model variations ignoring the re-scaled weights, and also followed Vandendijck et al. (2016)³⁴ by including the re-scaled weights in B-spline basis functions³⁷ in the predictor for HIV positivity. We chose B-splines having 3 degrees of freedom. We fitted eight possible model variations excluding and including country-level random effects cluster-level random effects and B-spline basis functions of the rescaled weights. Models were compared using the Watanabe-Akaike Information Criterion (WAIC)³⁸ and area under the receiver-operating characteristic (ROC) curve³⁹ for HIV status based on 10-fold cross-validation^{40,41} from each model (see Supplemental Digital Content 1).

Analysis

We began with a search for plausible interaction effects among the predictors using interaction plots. Pairs of predictors which showed evidence of non-parallel responses with HIV status were further examined using preliminary frequentist non-spatial mixed effects logistic models wherein TSNU was the sole random effect. Interactions for which $P < 0.05$ from the frequentist models were included, along with their main effects, in a preliminary full joint Bayesian spatial model which included country- and cluster-level random effects and B-spline basis functions of the rescaled weights. That full model was refitted sequentially eliminating unimportant interaction terms. The fixed-effect risk covariates and their interactions from the preliminary model having the smallest WAIC were retained for final fitting of the eight joint spatial models obtained by including/excluding country- and cluster-level random effects and the smoothed rescaled survey weights.

A detailed model description, including specifications of priors, hyperpriors and example model code, is given in Supplemental Digital Content 1. All computations were performed using R version 4.3.1⁴². The joint spatial model variations were fitted using the Integrated Nested Laplace Approximation (INLA) package⁴³, version INLA_2023-09-09.

Results

Interaction effects between covariates

Graphical analyses and preliminary frequentist models suggested the plausible existence of interaction effects between age and age of sexual debut, age group and partner HIV status, age of sexual debut and history of work during the past year, educational attainment and partner HIV status, and history of pregnancy and partner HIV status in the prediction of HIV status of AGYW (see Supplemental Digital Content 1). Therefore, we fitted preliminary full models, including those interaction effects, country- and cluster-level random effects and containing the smoothed rescaled survey weights to confirm the importance of any of those interactions. The age $\times$ partner HIV-status and education $\times$ partner HIV-status interactions were eliminated based on WAIC differences ≥ 10 between the full and reduced models.

Final set of behavioral and demographic risk factors

Our final set of behavioral/demographic risk factors were age group, urban versus rural residency, educational attainment, history of pregnancy, partner HIV status, transactional sex during the past year, age category at sexual debut, employment during the past year, and age group $\times$ age at sexual debut, age group $\times$ employment, and partner HIV status $\times$ history of pregnancy interactions. Those risk factors were included in the eight joint spatial models defined by the inclusion or exclusion of the country-level and cluster-level random effects, and the B-spline basis function of the re-scaled survey weights for prediction of HIV status. The areas under the ROC curves (AUC) were similar across all models which included the country-level random effect (Table 2). Among those, models M100, M101 and M111 produced WAIC values which were equal for practical purposes. Therefore, we chose the simplest of those, M100, including the country-level random effects but excluding the cluster-level random effects and the smoothed weights, as our final model. The rescaled weights conveyed almost no information about HIV test positivity (Supplemental Digital Content Fig. 5).

Predictors of HIV positivity: Population level viral load

Area-level PVL was—by far—the most important predictor of HIV status among the AGYW. The adjusted odds ratio (aOR) for the effect of PVL (70.6; 0.95-probability credible interval 20.7–240.5) represents the change in $\text{logit}(\text{positivity})$ for each $\frac{1}{2}$ -unit change in PVL (Table 3). The distribution of area-level PVL estimates was slightly skewed (Supplemental Digital Content 1) and ranged from 0–0.5449 $\log_{10}(\text{RNA copies}\cdot\text{ml}^{-1} + 1)$. The combination of joint estimation and spatial smoothing shrunk the highly unstable survey domain estimates of PVL, which ranged from 0–1.06 $\log_{10}(\text{RNA copies}\cdot\text{ml}^{-1} + 1)$ units, toward the mean as expected (Supplemental Digital Content 1).

Approximately 21% of the among-TSNU variation in PVL is attributable to spatial correlation, whereas essentially all the among-TSNU variation in HIV positivity is spatially independent (Table 3). Evidently the combination of fixed- and random-effect predictors, including PVL, exhausts the possibility for among-TSNU spatial correlation in the prediction of HIV positivity in this area-level model.

Strong spatial patterns in PVL are evident at both the subcontinental and national scales. Areas of higher PVL are more prevalent within the southern African countries of Malawi, Zambia, Zimbabwe, Lesotho and Eswatini than elsewhere (Fig. 1, Top). Visualization of PVL and the numbers of HIV-negative AGYW within countries provides a basis for aligning HIV prevention interventions to the geographic areas which might produce the greatest benefits in terms of infections averted (Fig. 1, bottom). Additional country-specific maps of PVL and numbers of at-risk AGYW are provided in Supplemental Digital Content 1.

Predictors of HIV positivity: Partner HIV Status and History of Pregnancy

Those AGYW who had both a partner with HIV and a history of pregnancy had the next strongest association with HIV positivity (aOR 32.06; 23.67–43.43), followed by those having a partner with HIV but no past pregnancy (aOR 12.12; 7.48–19.63) (Table 3). AGYW who reported having no partner but having a history of pregnancy, and AGYW

who had a partner of unknown HIV status regardless of history of pregnancy also incurred increased odds of HIV infection. There was no detectable association between history of pregnancy and HIV infection among AGYW who had HIV-negative partners.

Predictors of HIV positivity: Education Status

Completion of secondary education (aOR 0.68; 0.54–0.87) and attainment of some post-secondary education (aOR 0.35; 0.24–0.51) were protective (Table 3). Having engaged in transactional sex within the past year was associated with HIV acquisition (aOR 1.27; 1.08–1.49). AGYW of ages 15–19 years who had their sexual debut at age 16 years or younger incurred increased odds of HIV acquisition (aOR 1.42; 1.10–1.84), as did those of age 20–24 years who had their sexual debut later than age 16 (aOR 2.17; 1.72–2.74) or at 16 or younger (aOR 2.94; 2.33–3.70).

ROC Curves

The area under the subcontinental cross-validated ROC curve was 0.82 (95%CI 0.81–0.83). Country-specific ROC curves varied; their AUC values ranged from 68.0–81.7% (Supplemental Digital Content 1). The model enables statistically principled classification of newly encountered AGYW with known sensitivity and specificity (Table 4). For example, enrollment in prevention interventions of all those HIV-negative AGYW who have a probability of test positivity of at least 0.0276 would select those for whom their risk profiles would predict positivity with 80% sensitivity and 66.1% specificity. The threshold probabilities of test positivity are easily computed from responses to the risk screening questions, country and TSNU of residence, the parameter estimates and country and TSNU random effect estimates (Supplemental Digital Content 1).

Discussion

HIV prevention programs in SSA may benefit from identification and enrollment of AGYW who are at high risk of HIV acquisition and thereby strengthen their resilience against HIV infection^{1,2,5,15–18} to reach epidemic control. Previous studies and initiatives have explored the use of such models to predict HIV acquisition among AGYW^{2,15–18}. However, our model is the first to incorporate covariates at the individual, relationship and environmental levels, and produced the largest cross-validated AUC among those studies reporting that measure.

The inclusion of the environmental exposure PVL, a proxy of the amount of viral exposure in a geographic area, was an overwhelmingly important predictor of HIV positivity among AGYW and strengthened the predictive capability of our model. This strong association between area-level PVL and HIV-positivity among AGYW implies the importance of viral exposure in a geographic community. Strong spatial patterns in PVL are evident and visualization of PVL, along with the numbers of HIV-negative AGYW within countries, provides a basis for aligning HIV prevention interventions to high risk geographic areas that might produce the greatest benefits in terms of infections averted.

Other partner-relationship-level factors helped improve the prediction of HIV acquisition among AGYW. Viral exposure from a current partner (e.g., spouse or boyfriend) with HIV

had the second strongest association with HIV acquisition in our model. An AGYW being of age 20–24 years and having a male partner, regardless of whether he is cohabitating or is HIV positive, increased risk of HIV acquisition over younger AGYW who reported not having a partner. AGYW of age 20–24 years who had early sexual debut at age 16 or less, or engaged in transactional sex, or had early pregnancy were at higher risk of HIV acquisition than younger AGYW who did not have such experiences. These relationship-level factors, along with the PVL environmental exposure, are vital to identifying HIV risk among AGYW.

Our model also showed that achieving secondary education was associated with decreased risk of HIV positivity. School attendance is inversely associated with engaging in risky sexual behavior among AGYW in SSA countries like South Africa⁴⁴. Rudgard and colleagues⁴⁵ showed that a lack of education is positively associated with HIV risk and poor HIV outcomes. Further, schools in and of themselves provide education to increase awareness of safe sexual behaviors while also strengthening peer support to help ensure safe sexual practices. One challenge in SSA has been to retain AGYW in secondary education, which is partly why they remain at higher risk of HIV when compared to their male counterparts. AGYW of school age who are caregivers or spouses are particularly vulnerable to dropping out of school because of competing demands on their time from their families⁴⁶. Lack of funds to pay school fees and access to schools and educational facilities have also been barriers to completing a secondary education in SSA⁴⁶.

The key strength of this study is that we used data from 13 countries to identify the most robust, cross-county risk predictors at the individual, interpersonal and environmental levels. The primary limitations of this study are the array of risk predictors available in the PHIA surveys and the temporally static nature of the model. We did not have access to consistent information about sex-based violence among AGYW, although the predictive capability of such violence has been mixed^{12,42}. Use of our model to prioritize AGYW for enrollment in prevention interventions is limited to those who can adopt our risk-screening covariates within the 13 countries included in this study. The model fit is based on single PHIA surveys from each country and therefore cannot predict temporal changes. Like all similar efforts, the accuracy of voluntary responses to risk questions limits performance. Adolescence and early adulthood are associated with high vulnerability and risk⁴⁷. Stigma related to adolescent sexuality and HIV risk may impede the sharing accurate responses to the screening questions about sexual behavior or factors associated with risk⁴⁸.

Despite these limitations, the findings from this study can both help to inform the placement of HIV prevention interventions based on PVL, and to prioritize individual AGYW who are at highest risk of acquiring HIV for enrollment. For the latter, a screening threshold based on the predicted probability of infection could be chosen based on the ROC curve results from this study. Then, a web- or tablet-based application could be developed to insert responses to the questions used in this study and the country- and TSNU-level random effects into the predictor equation to estimate the probability of HIV infection among newly encountered AGYW. AGYW for whom their predicted probability of infection meets or exceeds the chosen threshold could be prioritized for enrollment in prevention interventions or additional services once enrolled. Our model indicates that, aside from age and area of

residence, as few as three and at most eight screening questions can classify risk among newly encountered AGYW with known sensitivity and specificity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This modeling study has been supported by the United States President's Emergency Plan for AIDS Relief (PEPFAR) through the U.S. Centers for Disease Control and Prevention (CDC). We are grateful to the many contributors to development and implementation of the PHIA surveys in Cameroon, Cote d'Ivoire, Eswatini, Ethiopia, Kenya, Lesotho, Malawi, Namibia, Rwanda, Tanzania, Uganda, Zambia and Zimbabwe. The findings and conclusions in this publication are those of the authors and do not necessarily represent the official position of the funding agencies.

Funding:

This publication and the PHIA surveys were funded by the President's Emergency Plan for AIDS Relief (PEPFAR). Technical assistance to the PHIA surveys was provided by the U.S. Centers for Disease Control and Prevention. Those surveys were conducted through cooperative agreements with grantees/federal entities, including country Ministries of Health and ICAP at Columbia University (New York, New York, USA). The PHIA surveys were reviewed and approved by the CDC and the institutional review boards of each partner implementing the surveys. The funder had no role in the design of this modeling study, analysis, decision to publish, or preparation of the manuscript.

Disclaimer:

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the funding agencies.

Data Availability

This study made secondary use of household survey data which are publicly accessible by reasonable request through the Population-Based HIV Impact Assessment website at <https://phia-data.icap.columbia.edu/datasets> (Accessed 2024-11-08). Naomi model estimates are publicly available at <https://naomi-spectrum.unaids.org/> (Accessed 2024-11-08). The 2023 UNAIDS/PEPFAR Combined Geographic Health Boundary file is available at https://adr.unaids.org/dataset/021426a1-c4f7-4fbd-82a0-36be60e71d1a/resource/796b9d0f-8348-4d80-af6c-d56bd370859b/download/2023_ssa_all_levels.geojson (Accessed 2024-11-08).

References

1. UNAIDS. World AIDS Day 2023 Fact Sheet: Global HIV Statistics. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2023. Accessed 2024-11-08. https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf
2. Rosenberg NE, Kudowa E, Price JT, et al. Identifying adolescent girls and young women at high risk for HIV acquisition: a risk assessment tool from the Girl Power-Malawi study. *Journal of Sexually Transmitted Diseases*. 2020;47(11):760–766. doi:10.1097/olq.0000000000001242 [PubMed: 33045165]
3. Low A, Thin K, Davia S, et al. Correlates of {HIV} infection in adolescent girls and young women in {Lesotho}: results from a population-based survey. *The Lancet HIV*. 2019;6(9):e613–e622. doi:10.1016/S2352-3018(19)30183-3 [PubMed: 31422056]

4. Price JT, Rosenberg NE, Vansia D, et al. Predictors of HIV, HIV risk perception, and HIV worry among adolescent girls and young women in Lilongwe, Malawi. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. 2018;77(1):53–63. doi:10.1097/QAI.0000000000001567 [PubMed: 28991885]
5. UNAIDS. HIV PREVENTION 2025 – ROAD MAP: Getting on track to end AIDS as a public health threat by 2030. 2022. Accessed 2024-11-08. <https://www.unaids.org/en/resources/documents/2022/prevention-2025-roadmap>
6. Fleischman J. Five Years of DREAMS and What Lies Ahead. 2021. <https://www.csis.org/analysis/five-years-dreams-and-what-lies-ahead>
7. USAID. Orphans and Vulnerable Children. United States Agency for International Development (USAID). Accessed 2024-04-26, <https://www.usaid.gov/global-health/health-areas/hiv-and-aids/technical-areas/orphans-and-vulnerable-children>
8. Birdthistle I, Kwaro D, Shahmanesh M, et al. Evaluating the impact of DREAMS on HIV incidence among adolescent girls and young women: A population-based cohort study in Kenya and South Africa. *PLoS Medicine*. 2021;18(10):e1003837. doi:10.1371/journal.pmed.1003837
9. Chabata ST, Hensen B, Chiyaka T, et al. The impact of the DREAMS partnership on HIV incidence among young women who sell sex in two Zimbabwean cities: results of a non-randomised study. *BMJ Global Health*. 2021;6(4):e003892. doi:10.1136/bmjgh-2020-003892
10. Pelletier A, Derado J, Maoela L, Lekhotsa T, Sechache M, Nkuatsana K. Impact of the DEAMS program on new HIV diagnoses in adolescent girls and young women attending antenatal care — Lesotho, 2015–2020. *Morbidity and Mortality Weekly Report*. 2022;71(2):48–51. doi:10.15585/mmwr.mm7102a3
11. Saul J, Cooney C, Hosseini PR, et al. Modeling DREAMS impact: trends in new HIV diagnoses among women attending antenatal care clinics in DREAMS countries. *journal article. AIDS*. 2022;36(Suppl. 1):S51–S59. doi:10.1097/qad.0000000000003259 [PubMed: 35766575]
12. Howard AL, Chiang L, Picchetti V, et al. Population estimates of {HIV} risk factors to inform {HIV} prevention programming for adolescent girls and young women. *AIDS Education and Prevention*. 2023;35(Supplement A):20–38. doi:10.1521/aeap.2023.35.suppA.20 [PubMed: 37406147]
13. UNICEF. Assessing the Vulnerability and Risks of Adolescent Girls and Young Women in Eastern and Southern Africa: A Review of the Tools in Use. United Nations Children’s Fund (UNICEF); 2021. Accessed 2024-11-08. <https://www.unicef.org/esa/media/9146/file/UNICEF-ESARO-AGYW-RV-Assessment-2021.pdf>
14. Howes A, Risher KA, Nguyen VK, et al. Spatio-temporal estimates of HIV risk group proportions for adolescent girls and young women across 13 priority countries in sub-Saharan Africa. *journal article. PLoS Global Public Health*. 2023;3(4):e0001731. doi:10.1371/journal.pgph.0001731
15. Brugh KN, Lewis Q, Haddad C, Kumaresan J, Essam T, Li MS. Characterizing and mapping the spatial variability of HIV risk among adolescent girls and young women: A cross-county analysis of population-based surveys in Eswatini, Haiti, and Mozambique. *journal article. PLoS ONE*. 2021;16(12):e0261520. doi:10.1371/journal.pone.0261520
16. Giovenco D, Pettifor A, MacPhail C, et al. Assessing risk for HIV infection among adolescent girls in South Africa: an evaluation of the VOICE risk score (HPTN 068). *Journal of the International AIDS Society*. 2019;22(7):e25359. doi:10.1002/jia2.25359
17. Moyo RC, Govindasamy D, Manda SO, Nyasulu PS. A prediction risk score for HIV among adolescent girls and young women in South Africa: identifying those in need of HIV pre-exposure prophylaxis. *HIV Research & Clinical Practice*. 2023;24(1):2221377. doi:10.1080/25787489.2023.2221377
18. Stoner M, Westreich D, Ahern J, et al. Modeling combination interventions to prevent human immunodeficiency virus in adolescent girls and young women in South Africa (HIV Prevention Trials Network 068). *Clinical Infectious Diseases*. 2021;73(7):e1911–e1918. doi:10.1093/cid/ciaa1598 [PubMed: 33325509]
19. Baral S, Logie CH, Grosso A, Wirtz AL, Beyrer C. Modified social ecological model: a tool to guide the assessment of the risks and risk contexts of HIV epidemics. *BMC Public Health*. 2013;13(1):482. doi:10.1186/1471-2458-13-482 [PubMed: 23679953]

20. PHIA. Population-Based HIV Impact Assessment. ICAP at Columbia University. Accessed 2024-11-08, 2023. <https://phia.icap.columbia.edu/>
21. UNAIDS. HIV sub-national estimates viewer. Joint United Nations Programme on HIV/AIDS (UNAIDS). Accessed 2024-11-08, <https://naomi-spectrum.unaids.org/>
22. 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. *Journal of the International AIDS Society*. 2021;24(S5):e25788. doi:10.1002/jia2.25788
23. Saul J, Bachman G, Allen S, F. Toiv N, Cooney C, Beamon TA. The DREAMS core package of interventions: A comprehensive approach to preventing HIV among adolescent girls and young women. *journal article*. PLoS ONE. 2018;13(12):e0208167. doi:10.1371/journal.pone.0208167
24. Fuller WA. *Measurement Error Models*. John Wiley & Sons; 1987.
25. Muff S, Riebler A, Held L, Rue H, Saner P. Bayesian analysis of measurement error models using integrated nested Laplace approximations. *journal article*. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 2015;64(2):231–252. doi:10.1111/rssc.12069
26. Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. Penalising model component complexity: A principled, practical approach to constructing priors. *journal article*. *Statistical Science*. 2017;32(1):1–28. doi:10.1214/16-sts576
27. Riebler A, Sørbye SH, Simpson D, Rue H. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *journal article*. *Statistical Methods in Medical Research*. 2016;25(4):1145–1165. doi:10.1177/0962280216660421 [PubMed: 27566770]
28. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *journal article*. *Annals of the Institute of Statistical Mathematics*. 1991;43(1):1–20.
29. Sørbye SH, Rue H. Scaling intrinsic Gaussian Markov random field priors in spatial modelling. *journal article*. *Spatial Statistics*. 2014;8:39–51. doi:10.1016/j.spasta.2013.06.004
30. Gelman A Struggles with survey weighting and regression modeling. *journal article*. *Statistical Science*. 2007;22:153–164. doi:10.1214/088342306000000691
31. Little RJ. To Model or Not To Model? Competing Modes of Inference for Finite Population Sampling. *journal article*. *Journal of the American Statistical Association*. 2004;99(466):546–556. doi:10.1198/016214504000000467
32. Paige J, Fuglstad G-A, Riebler A, Wakefield J. Design- and model-based approaches to small-area estimation in a low- and middle-income country context: Comparisons and recommendations. *journal article*. *Journal of Statistics and Survey Methodology*. 2022;10(1):50–80. doi:10.1093/jssam/smaa011
33. Wakefield J, Okonek T, Pedersen J. Small area estimation for disease prevalence mapping. *journal article*. *International Statistical Review*. 2020;doi:10.1111/insr.12400
34. Vandendijck Y, Faes C, Hens N. Prevalence and trend estimation from observational data with highly variable post-stratification weights. *journal article*. *Annals of Applied Statistics*. 2016;10(1):94–17. doi:10.1214/15-aos874
35. Zheng H, Little RJA. Penalized spline model-based estimation of the finite populations total from probability-proportional-to-size samples. *journal article*. *Journal of Official Statistics*. 2003;19(2):99–117.
36. Pfeffermann D, Skinner CJ, Holmes DJ, Goldstein H, Rasbash J. Weighting for unequal selection probabilities in multilevel models. *journal article*. *Journal of the Royal Statistical Society Series B*. 1998;60(1):23–40. doi:10.1111/1467-9868.00106
37. Eilers PHC, Marx BD. Flexible smoothing with (B)-splines and penalties. *journal article*. *Statistical Science*. 1996;11(2):89–121. doi:10.1214/ss/1038425655
38. Watanabe S A widely applicable Bayesian information criterion. *Journal of Machine Learning Research*. 2013;14:867–897.
39. Pepe MS. Receiver operating characteristic methodology. *Journal of the American Statistical Association*. 2000;95(499):308–311. doi:10.1080/01621459.2000.10473930
40. Kim J-H. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics and Data Analysis*. 2009;53(11):3735–3745. doi:10.1016/j.csda.2009.04.009

41. Rodríguez JD, Pérez A, Lozano JA. *k*-fold cross validation, prediction error, error estimation, bias and variance, decomposition of the variance, sources of sensitivity, supervised classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2010;32(3):569–575. doi:10.1109/TPAMI.2009.187 [PubMed: 20075479]
42. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2023.
43. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B*. 2009;71(2):319–392. doi:DOI: 10.1111/j.1467-9868.2008.00700.x
44. George G, Beckett S, Reddy T, et al. The role of schooling and comprehensive sexuality education in reducing HIV and pregnancy among adolescents in South Africa. *Journal of Acquired Immune Deficiency Syndromes*. 2022;90(3):270–275. doi:10.1097/QAI.0000000000002951 [PubMed: 35302969]
45. Rudgard W, Saminathen M, Banounin B, Shenderovich Y, Toska E. The role of structural factors for preventing HIV risk practices among adolescents in South Africa: A three-wave analysis of caregiving, education, food security, and social protection. 2023.
46. Psaki S, Haberland N, Mensch B, Woyczynski L, Chuang E. Policies and interventions to remove gender-related barriers to girls' school participation and learning in low- and middle-income countries: A systematic review of the evidence. *Campbell Systematic Reviews*. 2022;18(1):e1207. doi:10.1002/cl2.1207 [PubMed: 36913193]
47. Faden VB, Day NL, Windle M, et al. Collecting longitudinal data through childhood, adolescence, and young adulthood: methodological challenges. *Alcohol Clin Exp Res*. Feb 2004;28(2):330–40. doi:10.1097/01.alc.0000113411.33088.fe [PubMed: 15112941]
48. Currie DW, Apondi R, West CA, et al. A comparison of two population-based household surveys in Uganda for assessment of violence against youth. *PLoS ONE*. 2021;16(12):e0260986. doi:10.1371/journal.pone.0260986

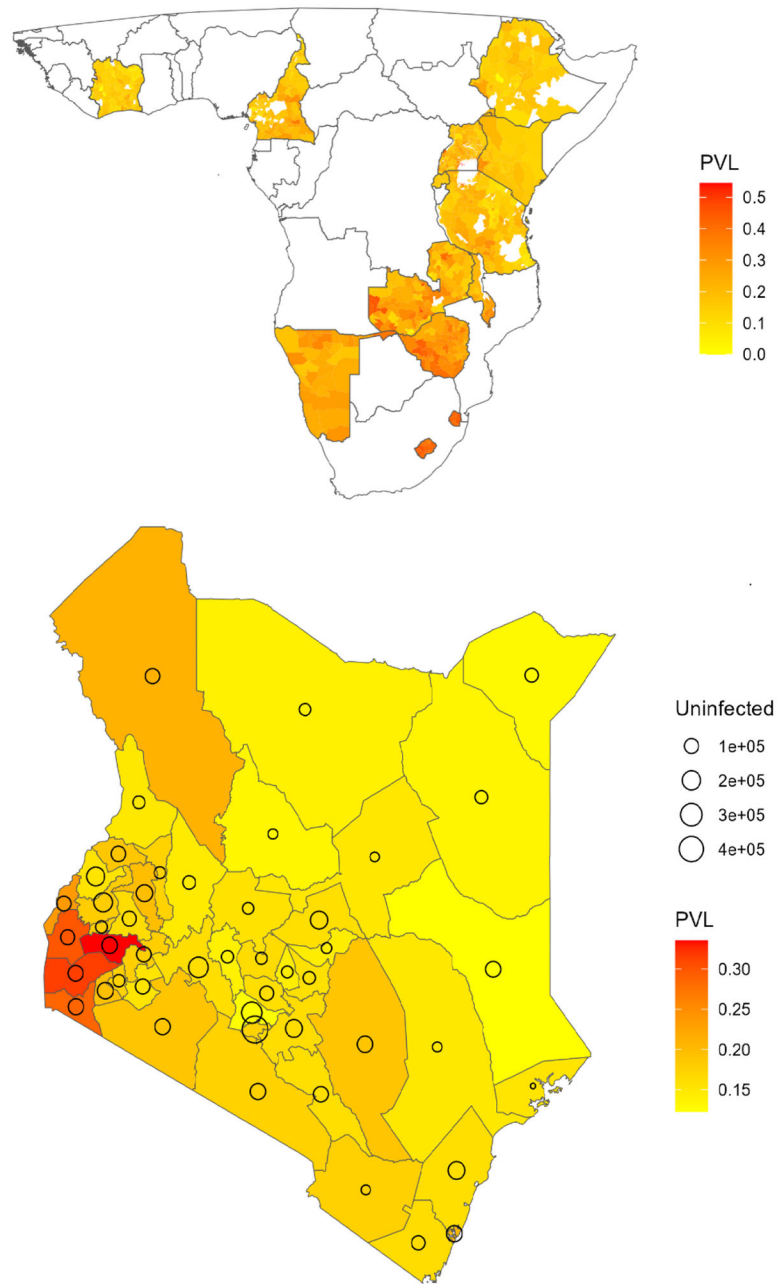


Figure 1.

Spatial distribution of HIV population viral load, PVL, at the subcontinental scale (top) and within Kenya (bottom). The estimates of the numbers of uninfected adolescent girls and young women during 2022 (bottom) were obtained from the Naomi model²².

Table 1.

Sample counts from 13 sub-Saharan African PHIA¹ surveys.

Country	Year(s)	Sampling strata		TSNU ²		Sampling clusters	Total respondents	HIV viral load tests ³	AGYW ⁴			Total
		Stratification	N	Level	name				15–19	20–24		
Cameroon	2017–18	Region × urbanicity ⁵	12	District	Sanitaire	168	481	980	2,550	2,242		4,792
Cote d'Ivoire	2017–18	District	14	District	Sanitaire	109	438	443	1,601	1,421		3,022
Eswatini	2016–17	Region × urbanicity	8	Region		4	286	3,003	1,010	856		1,866
Ethiopia	2017–18	Region (urban only)	11	Zone		109	393	19,136	2,019	1,567		3,586
Kenya	2018	County × urbanicity	92	County		47	798	27,745	2,363	2,130		4,493
Lesotho	2016–17	District	10	District		10	418	11,682	1,133	1,152		2,285
Malawi	2015–16	Health Zone	7	District		27	498	17,187	1,616	1,890		3,506
Namibia	2017	Region	14	District		38	460	16,939	1,432	1,389		2,821
Rwanda	2018–19	Province	5	District		30	373	30,637	3,322	2,707		6,029
Tanzania	2016–17	Region	31	District		178	521	31,579	2,928	2,772		5,700
Uganda	2016–17	Region	10	District		146	520	29,024	3,232	2,999		6,231
Zambia	2016	Province	10	District		115	511	19,115	2,413	2,025		3,972
Zimbabwe	2015–16	Province	10	District		63	499	22,501	2,091	1,777		3,868
Totals						1,044	6,196	280,323	24,877	27,322		52,171

¹ <https://phia.icap.columbia.edu/>

² Target subnational unit level for spatial smoothing and estimation of population viral load, typically the second subnational level.

³ Among total respondents

⁴ Adolescent girls and young women by age group (years)

⁵ Urbanicity is urban versus rural population density

Table 2.

Model comparison based on the Watanabe-Akaike Information Criterion (WAIC; smaller is better) and area under the cross-validated receiver-operating characteristic curves (AUC; larger is better) for candidate Bayesian joint spatial models of probability of HIV acquisition by adolescent girls and young women across 13 sub-Saharan African countries, 2015–2019.

Model	Random effects			WAIC	AUC (95% CL)	
	Country	Cluster	Weighting [‡]			
M000	–	–	–	703598	81.75	(80.78–82.72)
M001	–	–	+	703601	81.48	(80.50–82.47)
M010	–	+	–	703571	81.63	(80.66–82.61)
M011	–	+	+	703599	81.68	(80.71–82.65)
M100	+	–	–	703552	82.05	(81.10–83.01)
M101	+	–	+	703551	82.04	(81.09–83.00)
M110	+	+	–	703560	82.04	(81.09–83.00)
M111	+	+	+	703548	82.05	(81.19–83.01)

[‡]Inclusion of rescaled survey weights as smoothed *B*-spline basis functions (see Supplement.

Table 3.

Parameter estimates from the final joint spatial model of risk of HIV acquisition by adolescent girls and young women from 13 sub-Saharan African countries, 2015–2019.

Parameter	Level of categorical predictors	Pr(0.95) credible limits ¹		
		Posterior mean	lowerupper	
Adjusted odds ratios				
Intercept for HIV positivity ²		0.00	0.00	0.00
Intercept for PVL ²		1.26	1.25	1.26
PVL (effect per 1/2-unit change in log ₁₀ (PVL + 1)) ³		70.60	20.70	240.53
Urbanicity (baseline: rural residency)	Rural (baseline)			
	Urban	1.65	1.48	1.85
Partner HIV status and history of pregnancy	Baseline: No partner, never pregnant			
	No partner, ever pregnant	1.78	1.34	2.36
	Partner HIV-negative, never pregnant	0.87	0.64	1.17
	Partner HIV-negative, ever pregnant	0.79	0.62	1.02
	Partner status unknown, never pregnant	1.43	1.10	1.86
	Partner status unknown, ever pregnant	1.98	1.55	2.53
	Partner HIV-positive, never pregnant	12.12	7.48	19.63
	Partner HIV-positive, ever pregnant	32.06	23.67	43.43
Educational attainment	Less than primary completion (baseline)			
	Completed primary	0.95	0.75	1.21
	Completed secondary	0.68	0.54	0.87
	Some post-secondary	0.35	0.24	0.51
Transactional sex in past yr	No (baseline)			
	Yes	1.27	1.08	1.49
Sexual debut, age, and employment during past year	Never had sex, age 15–19, no employment during past year (baseline)			
	Debut > 16, age 15–19, no employment during past year	1.00	0.72	1.39
	Debut ≤ 16, age 15–19, no employment during past year	1.42	1.10	1.84
	Never had sex, age 20–24	0.75	0.50	1.14
	Debut > 16, age 20–24	2.17	1.72	2.74
	Debut ≤ 16, age 20–24	2.94	2.33	3.70

Parameter	Level of categorical predictors	Pr(0.95) credible limits ¹		
		Posterior mean	lower	upper
<i>Random-effect estimates</i>				
Standard deviation of linear predictor for HIV positivity ¹	Never had sex, employed during past year	0.74	0.48	1.14
Standard deviation of country-level random effect ¹	Debut > 16, employed during past year	1.27	1.08	1.48
Standard deviation of BYM2 area-level random effect for prevalence ¹	Debut ≤ 16, employed during past year	1.30	1.08	1.55
Fraction spatial component in area-level random error for prevalence ¹		0.2187	0.2080	0.2297
Fraction spatial component in area-level random error for PVL ¹		0.0068	0.0000	0.0395
Fraction spatial component in area-level random error for PVL ¹		0.1149	0.0142	0.3610
Fraction spatial component in area-level random error for PVL ¹		0.2061	0.0031	0.8314

¹Given the model and data, the true value is within limits with 0.95 probability

²See Supplement for detailed explanation

³A ½ unit change in log10(PVL + 1) is roughly equivalent to the range of observation (0–0.55), so that this odds ratio is approximately comparable to those for the categorical predictors

Table 4.

Estimated out-of-sample sensitivities and specificities (95% bootstrap confidence limits) attainable by prioritizing those adolescent girls and young women for enrollment who have predicted probabilities testing HIV positive that equal or exceed the Pr(positive) threshold. Out-of-sample performance was estimated from 10-fold cross-validation.

Sensitivity (%)		Specificity (%)		Pr(positive) threshold
90.0	(88.7–91.4)	51.1	(47.2–54.5)	0.0172
85.0	(83.5–86.7)	59.7	(57.2–62.4)	0.0225
80.0	(78.2–81.7)	66.1	(64.5–68.4)	0.0276
75.0	(73.2–77.0)	71.0	(69.6–73.1)	0.0323
70.0	(68.1–72.0)	75.8	(74.1–78.2)	0.0380
65.0	(62.9–67.0)	81.1	(79.0–82.8)	0.0467
60.0	(57.9–62.1)	85.0	(83.3–87.0)	0.0553
55.0	(52.8–57.2)	88.6	(87.4–89.8)	0.0665
50.0	(47.8–52.2)	91.2	(90.0–92.3)	0.0776
45.0	(42.9–47.3)	93.4	(92.5–94.2)	0.0918
40.0	(37.8–42.2)	95.1	(94.4–95.9)	0.1090
35.0	(32.8–37.1)	96.5	(96.0–97.1)	0.1292
30.0	(27.8–32.1)	97.7	(97.2–98.0)	0.1557