Estimation of HIV prevalence and burden in Nigeria: a Bayesian predictive modelling study

Amobi Andrew Onovo,[a,*] Adedayo Adeyemi,[b] David Onime,[c] Michael Kalnoky,[d] Baboyma Kagniniwa,[e] Melaku Dessie,[e] Lana Lee,[e] Deidra Parrish,[e] Bashorun Adebobola,[f] Gregory Ashefor,[g] Otse Ogorry,[h] Rachel Goldstein,[c] and Helina Meri[c,i]

[a]Institute of Global Health, University of Geneva, Switzerland
[b]Center for Infectious Diseases Research and Evaluation, Lafia, Nigeria
[c]Office of HIV/AIDS and TB, USAID, Nigeria
[d]IBTCI, Global Health Technical Assistance and Mission Support Project (GH-TAMS), Washington, DC, United States
[e]United States Agency for International Development, Bureau of Global Health, Office of HIV/AIDS, Washington, DC, United States
[f]National AIDS and STDs Control Programme, Abuja, Nigeria
[g]National Agency for the Control of AIDS (NACA), Abuja, Nigeria
[h]Data.FI, Palladium, Abuja, Nigeria
[i]U.S. Army Medical Research Directorate – Africa, Nigeria

eClinicalMedicine
2023;62: 102098

Published Online xxx
https://doi.org/10.1016/j.eclinm.2023.102098

Summary

Background The cost of population-based surveys is high and obtaining funding for a national population-based survey may take several years, with follow-up surveys taking up to five years. Survey-based prevalence estimates are prone to bias owing to survey non-participation, as not all individuals eligible to participate in a survey may be reached, and some of those who are contacted do not consent to HIV testing. This study describes how Bayesian statistical modeling may be used to estimate HIV prevalence at the state level in a reliable and timely manner.

Methods We analysed national HIV testing services (HTS) data for Nigeria from October 1, 2020, to September 30, 2021, to derive state-level HIV seropositivity rates. We used a Bayesian linear model with normal prior distribution and Markov Chain Monte Carlo approach to estimate HIV state-level prevalence for the 36 states +1 FCT in Nigeria. Our outcome variable was the HIV seropositivity rates and we adjusted for demographic, economic, biological, and societal covariates collected from the 2018 Nigeria HIV/AIDS Indicator and Impact Survey (NAIIS), 2018 Nigeria Demographic and Health Survey (NDHS) and 2016-17 Multiple Indicator Cluster Surveys (MICS). The estimated population of 15–49 years olds in each state was multiplied by estimates from the estimated prevalence to generate state-level HIV burden.

Findings Our estimated national HIV prevalence was 2.1% (95% CI: 1.5–2.7%) among adults aged 15–49 years in Nigeria, which corresponds to approximately 2 million people living with HIV, compared to previous national HIV prevalence estimates of 1.4% from the 2018 NAIIS and UNAIDS estimation and projection package PLHIV estimation of 1.8 million in 2022. Our modelled HIV prevalence in Nigeria varies by state, with Benue (5.7%, 95% CI: 5.0–6.3) having the highest prevalence, followed by Rivers (5.2%, 95% CI: 4.6–5.8%), Akwa Ibom (3.5%, 95% CI: 2.9–4.1%), Edo (3.4%, 95% CI: 2.9–4.0%) and Taraba (3.0%, 95% CI: 2.6–3.7%) placing fourth and fifth, respectively. Jigawa had the lowest HIV prevalence (0.3%), which was consistent with prior estimates.

Interpretation This model provides a comprehensive and flexible use of evidence to estimate state-level HIV seroprevalence for Nigeria using program data and adjusting for explanatory variables. Thus, investment in program data for HIV surveillance will provide reliable estimates for HIV sub-national monitoring and improve planning and interventions for epidemiologic control.

Funding This article was made possible by the support of the American people through the United States Agency for International Development (USAID) under the U.S. President's Emergency Plan for AIDS Relief (PEPFAR).

Keywords: HIV prevalence; HIV burden; Bayesian model; Markov Chain Monte Carlo; Nigeria

*Corresponding author.
E-mail address: amobiandrewonovo@gmail.com (A.A. Onovo).

Research in context

Evidence before this study

The cost of doing a population-based survey in Nigeria is expensive and securing funding for a national population-based survey may not be guaranteed for several years. On March 26, 2022, we searched Google scholar, PubMed, JSTOR and the American Medical Association website for published research articles using the terms "HIV prevalence", "Population-based surveys", "Bayesian modelling", "Markov Chain Monte Carlo", "Burden of HIV estimation", and "HIV prevalence trends" with no language or date restrictions. Our search turned up descriptive, retrospective, systematic review, and meta-analysis papers, that indicate that Bayesian statistical models have been applied in a broad variety of epidemiological applications, and that robust HIV prevalence estimation models with multilevel correlation structures, multiple measurement scales for the different input variables, and nonlinear regression frameworks can all be fitted using a Bayesian Markov Chain Monte Carlo technique.

Added value of this study

The authors developed a flowchart of the Bayesian model designed to estimate new HIV prevalence in Nigeria, analyzed national HIV testing services (HTS) data to derive state-level HIV seropositivity rates, and utilized a Bayesian Generalized Linear Model based on the Markov Chain Monte Carlo approach to estimate new (2.1% as opposed to 1.4%) HIV state-level prevalence for the 36 states + 1 FCT in Nigeria. We additionally employed the exploratory data analysis (EDA) technique to uncover patterns, and association between Nigeria's estimated HIV burden and coronavirus disease. According to our study, HIV prevalence varies by geographic region, and since the last population-based survey in 2018, the HIV burden has grown by 7.2 percentage points (i.e., from 1,869,259 based on the UNAIDS 2022 spectrum estimates to 2,004,068 based on our Bayesian model). According to the findings at the 95% credible interval, HIV prevalence decreases by 0.53 (−6.36 to 5.04), 0.01 (−0.09 to 0.07), 0.05 (−0.17 to 0.08), and 0.05 (−0.16 to 0.06) percentage points per unit increase in the number of people who test negative for HIV, in ANC coverage among antenatal clients, in the wealth index of the poor and the proportion of literate women, respectively.

Implications of all the available evidence

Our study's estimated mean HIV prevalence was 1.5 times higher than the estimated mean HIV prevalence from the 2018 Nigeria AIDS Indicator and Impact Survey (NAIIS), implying an upward trend in HIV infection in the three years since the last NAIIS, and previous national HIV prevalence estimates for the 36 states + 1 FCT may have been underestimated in several cases. The use of a Bayesian statistical approach to combine multiple datasets into a quantitative framework has a lot of potential for HIV epidemiology model fitting, and it may be a cost-effective way that can be utilized to enhance integrated surveillance initiatives for infectious diseases response at national and sub-national levels in Nigeria.

Introduction

In Nigeria, 80% of new HIV infections are caused by unprotected heterosexual intercourse, with most remaining HIV infections happening in key populations such as sex workers, men who have sex with men, people who inject drugs and transgender people.[1–3] In West Africa, HIV prevalence is low, with adult prevalence in the general population estimated at 2% or less.[4] Adult HIV prevalence in Nigeria was estimated at 1.4% among people aged 15–49 years in the 2018 Nigeria AIDS indicator and Impact Survey (NAIIS). NAIIS was a population-based survey that was conducted to track key national HIV-related indicators, such as progress toward the UNAIDS 95-95-95 targets as well as to guide policy and funding priorities.[5] The President's Emergency Plan for AIDS Relief (PEPFAR) mainly provided funding, technical and logistic support for NAIIS implementation in Nigeria with some additional resources from the Global Fund.

Accurate estimates of HIV prevalence are crucial for tracking the HIV epidemic, planning, developing, implementing, and assessing preventive and treatment programs, and projecting resource demands.[6–9] National population-based surveys[10–12] have become a significant data source for estimating HIV seroprevalence in several countries in sub-Saharan Africa.[13,14] The fact that not all persons eligible to participate in a survey can be contacted and that some of those who are contacted do not consent to HIV testing might threaten the validity of survey-based prevalence estimates. This is in addition to the inability to reach the homeless or some adults staying in non-residential places during the survey. Incomplete HIV testing can contribute to selection bias. Interestingly, a recent study found evidence of significant downward bias in existing national HIV prevalence estimates for Zambian men due to selective survey non-participation.[15] An analysis of the Demographic and Health Surveys (DHS), which are the most common nationally representative surveys for HIV prevalence in Sub-Saharan Africa, reveals that 23% of adult men and 16% of adult women in the region do not participate in HIV testing, with as high as 37% for men in Zimbabwe in 2005–2006 and as low as 3% for women in Rwanda 2005.[16] The most recent national population-based study in South Africa, the total non-participation rate for HIV testing among adults was found to be 32%.[11] NAIIS had an overall non-participation or response rate of up to 11%.[17] Individuals who are HIV-positive (from previous testing) may be afraid of stigma, marginalization, or mistreatment if others learn of their HIV status.[18,19]

Individuals who suspect they may be HIV-positive (based on past sexual behavior, for example) may be afraid of having their concerns confirmed.[20] A longitudinal study in Malawi found that people who had previously tested HIV-positive were 4.6 times less likely than those who had previously tested HIV-negative to consent to a fresh HIV test.[21] Empirical findings like this help drive home the need to improve HIV prevalence and PLHIV burden estimation in Nigeria, as well as strengthen national and sub-national surveillance systems.

The Estimation and Projection Package (EPP) of the Joint United Nations Program on AIDS (UNAIDS) was created as a tool for constructing national and sub-national HIV epidemic curves. Except for countries such as South Africa and India, the EPP is presently utilised by other countries. Currently, Nigeria uses the UNAIDS Spectrum tool as the main estimation model to generate state-level PLHIV estimates. Every year, the UNAIDS secretariat collaborates with the Nigerian government, US government agencies, and other key stakeholders to construct PLHIV estimates based on Spectrum files created and maintained by a national epidemic surveillance team. The UNAIDS Spectrum tool incorporates assumptions from the Naomi tool and findings from NAIIS 2018. Estimates are based on data from the 2006 population census and programmatic data. The Spectrum files are complicated and inflexible, making it cumbersome to acquire results.[22] Over the years, some estimates of the PLHIV burden have been inconsistent and/or contradictory, resulting in skepticism/doubt about programmatic outcomes especially at the sub-national level. Overestimation or underestimation of PLHIV burden in some states of the country, and national percentage of coverage of prevention of mother-to-child transmission of HIV (PMTCT), which stakeholders found low for the country despite years of investments and achievements in PMTCT, are examples of such inconsistent results. Spectrum constraints are exacerbated by the lack of up-to-date national and sub-national population census data.

In this paper, to effectively estimate the HIV prevalence for Nigeria, we provide a new statistical approach for improving the efficiency of estimates to support HIV programming at national and sub-national levels. We incorporated sociodemographic, economic, behavioral, and biological risk factors in a Bayesian statistical model using the national program level HIV seropositivity data as the response variable in our model. We utilized Bayesian statistical model and estimated proportion of persons in a population who are living with HIV at a specific point in time. Consequently, we calculated the burden of HIV (i.e., the number of People Living with HIV/AIDS) by dividing the estimated HIV prevalence by 100 and multiplying by their respective state population census figures for the 15–49 years age group.

Methods

Study design and participants

In this cross-sectional study, we developed and validated a Bayesian statistical model that aimed to estimate new HIV prevalence for Nigeria and calculate HIV burden state-by-state. We gathered and synthesised data from a variety of sources, including national/PEPFAR program data between October 1, 2020 and September 30, 2021, as well as data from population-based surveys such as the NAIIS 2018, Nigeria Demographic and Health Survey, and the 2016/2017 Multiple Indicator Cluster Survey (Fig. 1). The data used in this study came from 36 + 1 FCT states encompassing Nigeria's six geopolitical zones, with an estimated total population of 206 million people at the mid-year in 2020, according to United Nations (UN) data.[23] The HIV seropositivity data from the PEPFAR program is the outcome variable. NAIIS 2018 prevalence, PEPFAR program data (.i.e., PMTCT coverage, ART coverage, currently on treatment, PMTCT negative, HTS negative, and viral load suppression), estimated 2020 population census (US Census Bureau[24]), fertility rate, teenage pregnancy, HIV knowledge, condom use, multiple partnership, ANC coverage, wealth index (poorest and richest), literacy rate (men and women), and male circumcision were among the 19 other variables considered as explanatory variables and were aggregated for adults aged 15–49 years (Table 1).

The differences in the model's covariates reflect the risk factors (i.e., demographic, biological, socioeconomic, and behavioral) that put people at risk of contracting and transmitting HIV. ART, according to several studies, reduces HIV transmission as well as illness and death among HIV-positive people. In Sub-Saharan Africa, ART is being expanded as a strategy for HIV prevention ("HIV treatment as prevention"). Our study focused on the 15–49-year-old age range, which is considered as the most sexually active age group in Nigeria. Compared to NAIIS HIV prevalence data which focused solely on the general population, many of those at highest risk, such as partners of people with HIV, young people in high HIV prevalence settings, and key populations (KP) are aggregated and included in the HIV seropositivity results from program data. This ensures full capability of using surveillance data, and provides the ability to include sub-national epidemics combining them with general population epidemics in our model to estimate the national HIV prevalence. The key populations (KP) group, which is at higher risk of HIV acquisition and transmission than the overall general population, was not specified in 2018 NAIIS. KP such as female sex workers, men who have sex with men, people who inject drugs, and transgender people are disproportionately affected by HIV. They account for almost 36% of the 1.9 million new adult HIV infections diagnosed each year.[25] In 2020, KP, their clients, and sexual

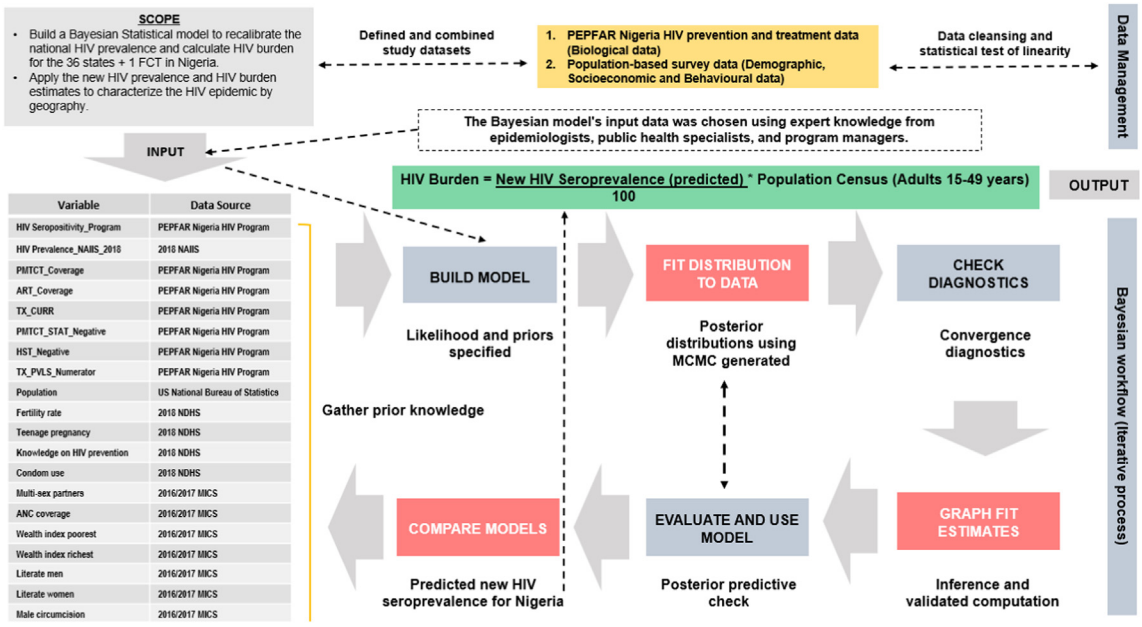

Fig. 1: **Study flowchart.** Several data sets from the national HIV program and population-based surveys were integrated. The Bayesian model was designed to estimate new HIV prevalence in Nigeria. For each of the 36 states +1 FCT in the model, a Bayesian generalised linear regression model based on the MCMC algorithm was employed to simulate 1000 iterations of HIV prevalence. The estimated HIV prevalence was used to calculate the new HIV burden for Nigeria.

partners accounted for 64% of new HIV infections in West and Central Africa, and 25% of new HIV infections in the East and Southern African subregion.[26] According to data from Senegal, Gambia, Cote d'Ivoire, Ghana, and Nigeria, a large proportion of HIV infections[27] occur among MSM, many of whom also report having intercourse with women.[28]

Statistical analysis

We created scatterplots to examine the relationship between all the predictor variables and HIV seroprevalence and utilized Pearson correlation coefficient to test the linearity assumption between each predictor variable and the outcome variable. We used the Bayesian linear model with normal prior distribution to estimate the proportion of persons in a population who are living with HIV as of 26 March 2022. Compared to the frequentist methods like OLS, Bayesian statistical models perform estimations using computational simulations based on the Markov-Chain Monte Carlo (MCMC) algorithms, and it generates 95% credible intervals, which have a more common-sense interpretation because credible intervals quantify the uncertainty estimated. There are three steps to our analysis: To begin, we fitted a Bayesian linear model specifying the Gaussian family distribution using Gibbs sampling, a multiparameter model of the Metropolis–Hastings algorithm based on the technique of MCMC to specify the posterior model. The parameters were then estimated iteratively until the burn-in conditions were met. In our study, the Normal likelihood function was specified. Simulates are then drawn from the posterior distribution to create an empirical distribution of likely values for the population parameter (Supplementary Material 4). The Gaussian distribution is known as one of the most important probability distributions in statistics because it can fit many natural phenomena, such as continuous variables like HIV prevalence in our study. The posterior model combines a probability distribution, which contains information about model parameters based on observed data, with a prior function that contains previous information about model parameters (before viewing the data). The model parameters included the response variable "HIV seropositivity from program's HTS data" and the independent covariates of interest; NAIIS 2018 prevalence, PMTCT coverage, ART coverage, currently on treatment, PMTCT negative, HTS negative, viral load suppression, estimated population census, fertility rate, teenage pregnancy, HIV knowledge, condom use, multiple partnership, ANC coverage, wealth index (poorest and richest), literacy rate (men and women), and male circumcision. Secondly, we computed Bayesian predictions for the outcome variable. Using the "posterior predict" function of the Bayesian statistical model in R, we estimated the proportion of persons in Nigeria who are living with HIV based on the distributions of the fitted posterior model. Here we simulated one thousand (1000) MCMC samples of outcome

Serial No.	Variable	Definition/Unit	Data type	Source	Year	Mean	SD	Range	Skewness
1	Population	Population estimates of all people (both sexes) between ages 15 and 49	Discrete	Estimated Nigeria population from US Census Bureau	2021	2,666,311	1,385,769	7,397,784	2.4
2	TX_CURR	Number of adults and children currently receiving antiretroviral therapy (ART)	Discrete	Program Data	2021	48,915	56,714	215,166	1.89
3	PMTCT_STAT_NEG	Number of women attending ANC1 who were tested for HIV and received a negative result.	Discrete	Program Data	2021	12,887	9670	33,120	0.79
4	HTS_TST_NEG	Number of individuals who received HIV Testing Services (HTS) and received a negative test result	Discrete	Program Data	2021	73,874	82,890	365,973	1.88
5	TX_PVLS_N	Number of ART patients with suppressed VL results (<1000 copies/ml) documented in the medical or laboratory records within the past 12 months	Discrete	Program Data	2021	39,785	47,229	191,608	2.03
6	PMTCT Coverage	Percentage of pregnant women attending ANC who were tested for HIV and received results	Continous	Program Data	2021	96.9	9.1	41.0	−3.01
7	ART Coverage	Percentage of people living with HIV receiving antiretroviral therapy	Continous	Program Data	2021	90.1	40.1	147.7	0.32
8	HIV Seropositivity Program	Percentage of positives found out of those who were tested and received their test results	Continous	Program Data	2021	2.0	1.4	6.4	1.53
9	NAIIS 2018	HIV prevalence among adults aged 15–49 years	Continous	NAIIS	2018	1.4	1.1	4.5	1.75
10	Fertility Rate	Total fertility rate for women aged 15–49 years	Continuous	NDHS	2018	5.1	1.1	3.9	0.41
11	Teenage Pregnancy	Percentage of women aged 15–19 who have begun childbearing	Continuous	NDHS	2018	16.7	9.5	39.6	0.52
12	Knowledge of HIV Prevention	Percentage of young women and young men aged 15–49 with comprehensive knowledge about HIV prevention	Continuous	NDHS	2018	13.5	5.7	25.6	0.3
13	Condom use among sexual partners	Percentage of women and men aged 15–49 who reported using condoms and limiting sexual intercourse to one uninfected partner	Continuous	NDHS	2018	75.0	15.1	52.2	−0.18
14	Multiple Sex partners	Among all men aged 15–49, percentage who had sexual intercourse with more than one sexual partner in the past 12 months	Continuous	NDHS	2018	12.8	6.2	32.4	1.13
15	ANC Coverage	Percentage of women aged 15–49 years who received antenatal care (ANC)	Continuous	MICS	2017	69.3	19.7	66.3	−0.41
16	Wealth index poorest	Percent distribution of the household population by poorest wealth index quintile	Continous	MICS	2017	14.7	17.3	50.3	0.82
17	Wealth index richest	Percent distribution of the household population by richest wealth index quintile	Continous	MICS	2017	23.8	18.7	80.8	1.12
18	Literate men	Number of men aged 15–49 years who attended secondary or higher education	Continous	MICS	2017	77.7	20.6	66.7	−0.7
19	Literate women	Number of women aged 15–49 years who attended secondary or higher education	Continous	MICS	2017	69.3	27.5	78.1	−0.42
20	Male circumcision	Number of men aged 15–49 years who report having been circumcised	Continous	MICS	2017	98.2	2.4	11.6	−2.7

The mean represents the dataset's average value. The standard deviation (SD) summarizes the differences between each observation from the mean, while the range represents the spread of study data from the lowest to the highest value in the distribution. The skewness matrix quantifies how far each variable's distribution deviates from the normal distribution. The distribution is highly skewed if skewness is less than −1 or greater than 1. The distribution is moderately skewed if the skewness is between −1 and −0.5 or between 0.5 and 1. The distribution is approximately symmetric or normally distributed if the skewness is between −0.5 and 0.5. Prior to model processing, the variables were normalized to a specific range.

Table 1: **Definition and descriptive statistics of variables in the Bayesian statistical model.**

values for each of the 36 + 1 FCT states and generated the posterior means and estimated p-values for each observation. We used a random-number seed to ensure reproducibility. Finally, we conducted posterior estimated checks by comparing the observed data with the MCMC replicates (simulated data from the posterior predictive distribution). Unlike classical prediction, which produces a single value for each observation, Bayesian prediction produces an MCMC sample of values for each observation.

We georeferenced the estimated HIV prevalence across 36 + 1 FCT states and generated a choropleth map.

Ethics

The analysis was conducted using routine data collected through the PEPFAR Nigeria HIV/AIDS program and survey data derived from the 2018 Nigeria Demographic and Health Survey and the 2013 Multiple Indicator Cluster Survey. In accordance with the Nigeria HIV program policy, informed consent was obtained from all clients who were tested for HIV. Only aggregate level data were analyzed in this study.

Role of the funding source

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. AAO, AA, DO had access to the dataset, and all the authors: AAO, AA, DO, MK, BK, MD, LL, DP, BA, GA, OO, RG, and HM accept responsibility for the decision to submit for publication.

Results

The flowchart for our study is detailed in Fig. 1. From the bivariate analysis, the scatterplot shows monotonic increase; this indicates a strong positive correlation between HIV seroprevalence from national program HTS data and 2018 NAIIS (r_s (36) = 0.73, p < 0.0001) (Fig. 2a). As the HIV prevalence from the 2018 NAIIS rises, the HIV prevalence from the national program rises as well. Individuals who were HIV-negative, virally suppressed, or who were currently on life-saving ART showed a strong positive association with the 2018 NAIIS (Fig. 2b). This finding highlights the need of integrating and utilizing program data for HIV surveillance, implying that HIV seropositivity from national program data can be adjusted or recalibrated to track and monitor HIV epidemics in the general population. Eleven variables of the nineteen variables included in our study showed a linear relationship with HIV seroprevalence in programs, and six of these variables (NAIIS 2018, TX_CURR, HTS negative, TX_PVLS, condom use by multiple sexual partners, ANC coverage) showed a positive linear relationship, while five variables (teenage pregnancy, fertility rate, poorest wealth index, and literate men and women) showed a negative linear relationship (Supplementary Material 6).

The summary of the fitted Bayesian linear model is provided in Table 1. The default priors used for the model parameters were normal (mean 0, standard deviation 10,000) for the regression coefficients and inverse gamma (shape 0.01, scale 0.01) for the variance parameter. In Table 1, the second and third columns of the Bayesian normal regression report the posterior means and standard deviations of the model parameters. The posterior mean estimates for the variance, 1.85, was close to the residual mean squared estimate, 1.70. The minimum efficiency in the model was 0.23, and the mean efficiency 0.89. Our acceptance rate (AR) was good, and efficiencies were high. We explored convergence by computing graphical diagnostic plots for all models to confirm this. Overall, graphical diagnostic plots showed that MCMC converged and mixes well for all parameters in the model (see graphical diagnostic plots in Supplementary Material 3).

The Bayesian regression analysis revealed an inverse association for seven of the model's variables: HIV prevalence from NAIIS 2018, HTS negative, population, fertility rate, ANC coverage, wealth index poorest and literate women (Table 2). According to the findings at the 95% credible interval, HIV prevalence decreases by 0.53 (–6.36 to 5.04), 0.01 (–0.09 to 0.07), 0.05 (–0.17 to 0.08), and 0.05 (–0.16 to 0.06) percentage points per unit increase in the number of people who test negative for HIV, in ANC coverage among antenatal clients, in the wealth index of the poor and the proportion of literate women, respectively. The Nigeria HIV/AIDS Indicator, and Impact Survey (NAIIS) provided the benchmark or baseline HIV prevalence data for characterizing the epidemic by geography, sex, and age to guide tailored public health response to the HIV epidemic across the community and health facility levels. In the NAIIS 2018 study, HIV prevalence was much higher in rural areas (≈2.0% vs. 1.4%) than in urban areas. Nigeria Key Population HIV program is a community-based program. To optimize HIV case finding and treatment interventions at local government levels, the KP program uses community-based structures such as community ART clinics and peer navigators. Key populations continue to be stigmatized and discriminated against, as well as arrested and criminalized, and are therefore driven underground and difficult to reach as a result. NAIIS 2018 study did not include key populations due to its sampling approach. Based on our study, HIV prevalence is predicted to decrease by 1.14 percent (–2.93 to 0.81) percentage points in the study regions and this reduction might be attributed to the combination of intensified prevention and treatment services targeted at both general populations and KPs. Nigeria has a mixed epidemic, which means that while HIV prevalence is high among the general population, certain groups continue to bear a far greater HIV burden than the rest of the population. The subpopulation group in Nigeria where HIV prevalence is still on the rise is men who have sex with men. In 2017, this group had a prevalence of 23%, which was significantly higher than the next highest prevalence group, sex workers, who had a prevalence of 14.4%.[29] According to recent UNODC studies on HIV prevalence in Nigerian prisons and drug use in Nigeria, 2.8% of inmates and 9% of people who inject drugs (PWIDs) are infected with HIV/AIDS.[30] These figures are significantly higher than the 1.4% national HIV prevalence

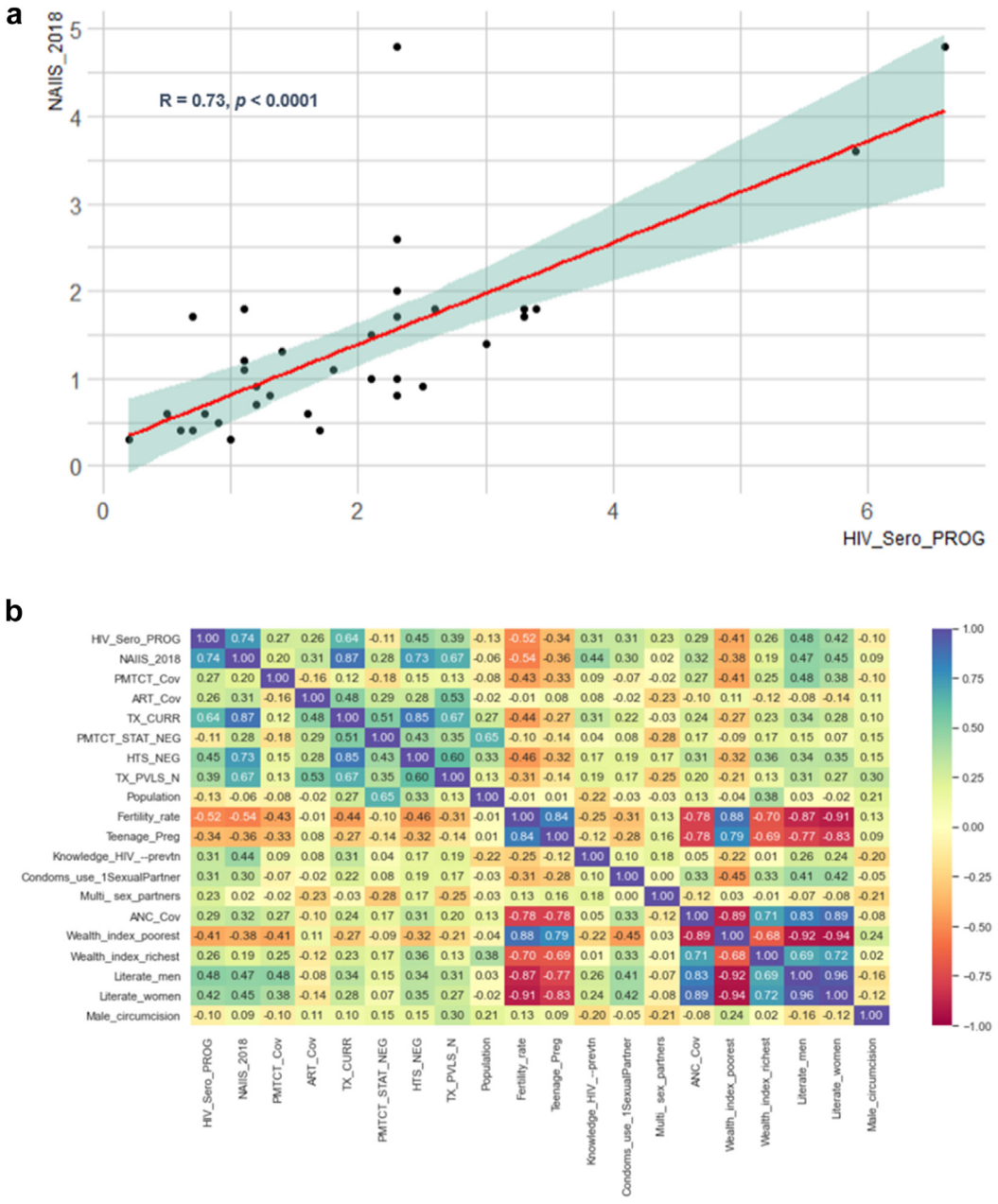

Fig. 2: Bivariate analysis of continuous variables in the Bayesian model (a) Scatter plot of HIV prevalence from 2018 NAIIS vs. HIV seropositivity from program's HTS data (b) Matrix plot showing correlation of Nineteen continuous variables and HIV seropositivity from program's data. The several black dots in (a) represent the 36 + 1 FCT states in Nigeria.

rate among the general population, implying the importance of incorporating and utilizing program data which includes both the general and priority populations. According to previous research,[31] HIV-positive women have decreased fertility due to fetal losses caused by HIV infection and co-infection with other sexually transmitted diseases. According to our findings, when fertility improves, HIV prevalence decreases by 0.54. Overall, our result implies that HIV prevalence estimates based on program data are responsive to population-level unit changes in demographic, biological, and behavioral variables. This also shows that when the quality of HIV sero-positivity data from national programs is improved, it can be used to track the HIV epidemic and evaluate effectiveness of HIV programs at the national and subnational levels without necessarily investing in expensive HIV surveys or non-routine data collection activities.

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred. Interval]	
HIV Prevalence (Program data)						
NAIIS_2018	−1.14758	0.932792	0.033065	−1.14147	−2.936072	0.810208
PMTCT_Cov	0.005301	0.041614	0.001409	0.005203	−0.083187	0.0846860
ART_Cov	0.001159	0.01004	0.000317	0.001262	−0.0187269	0.0209165
TX_CURR	5.331814	4.633695	0.174419	5.358043	−3.562352	14.192610
PMTCT_STAT_NEG	2.529787	2.060404	0.062728	2.609106	−1.873445	6.522597
HTS_NEG	−0.53045	2.935785	0.093681	−0.60135	−6.362088	5.0478930
TX_PVLS_N	2.494055	2.035624	0.06144	2.533955	−1.431750	6.473920
Population	−6.54407	2.768638	0.087552	−6.61846	−11.655030	−0.84663
Fertility_rate	−0.54035	0.744729	0.024191	−0.52138	−2.0258710	0.8525370
Teenage pregnancy	0.031475	0.057126	0.001982	0.033318	−0.0810859	0.1432969
Knowledge HIV prevention	0.006438	0.054991	0.001554	0.005067	−0.1015686	0.1167542
Condoms use_Sexual Partner	0.011067	0.022711	0.000718	0.01004	−0.0315753	0.0546242
Multi Sex partners	0.060319	0.064329	0.002034	0.06059	−0.0697243	0.182168
ANC_Cov	−0.00683	0.041087	0.001321	−0.00833	−0.0943591	0.0749489
Wealth_index_poorest	−0.04821	0.065952	0.002086	−0.04638	−0.1767972	0.08037
Wealth_index_richest	0.035477	0.027341	0.000911	0.037163	−0.0217615	0.0886151
Literate_men	0.041048	0.062038	0.001667	0.041518	−0.0822013	0.1596434
Literate_women	−0.05628	0.057989	0.001618	−0.05514	−0.1693237	0.0607161
Male_circumcision	0.027539	0.12198	0.004651	0.029896	−0.2167371	0.2498802
_cons	2.14474	12.72879	0.416765	1.901843	−22.45150	27.33074
var	1.859354	0.786299	0.051747	1.704222	0.8835566	4.161198

MCMC iterations = 3500; Burn-in = 2500; MCMC sample size = 1000; Number of obs = 37; Acceptance rate = 1; Efficiency: min = 0.230; avg = 0.8979; max = 1. Burn-in is the number of iterations thrown away at the start of the MCMC run, while MCMC iterations are random samples from the posterior means. The MCMC sample size is the number of MCMC draws used to calculate the Bayesian credible bounds. The acceptance rate is the percentage of simulations that were used, whereas the efficiency is the model's performance.

Table 2: **Bayesian normal regression using Gibbs sampling.**

Table 3 shows the estimated HIV prevalence and burden for Nigeria's 36 + 1 states. The midpoint values, lower credible, and upper credible intervals were used to represent the predicted prevalence and HIV burden. In our analysis, HIV prevalence ranged from 0.3% in Jigawa to 5.7% in Benue state. In comparison to the prevalence in the other states, Rivers (5.2%) and Akwa Ibom (3.5%) have high HIV prevalence. In Spectrum, the HIV infection pattern by state was comparable to our modelled estimates, with Benue (5.1%), Akwa Ibom (4.5%), and Rivers (3.9%) having higher HIV prevalence than the other states, and HIV prevalence being lowest in Jigawa state, as shown in both Spectrum (0.4%) and our study estimations (0.3%).

When the modelled mean HIV prevalence was compared to the mean Spectrum prevalence, HIV prevalence varied substantially among geographic regions, but was identical in the South East (NC: 2.74 vs. 1.67; NE: 1.50 vs. 1.13; NW: 1.04 vs. 0.58; SE: 1.94 vs. 1.86; SS: 3.3 vs. 2.7; and SW: 1.9 vs. 0.93). This could indicate that the HIV epidemic has shifted/increased geographically in the four years since the last population-based study in 2018 (Fig. 3a). Overall, the HIV burden from our study (Fig. 3b) suggests that it overlaps (HIV prevalence from Spectrum stayed within the modelled HIV prevalence's lower and upper credible intervals) with the HIV burden from Spectrum [North Central: 55,699 (47,542–72,217); NE: 27,026 (18,595–40,348); NW: 28,996 (18,515–53,076); SE: 58,614 (31,830–58,164); SS: 92,310 (75,948–106,104); and SW: 44,550 (46,401–85,591)]. In comparison to the other regions, HIV burden was highest in the SS (91,053) and lowest in the NE (28,781), while the pattern was comparable for HIV burden from Spectrum in the SS (92,310) and NE (27,026).

The study used a choropleth map in Fig. 4 to highlight the spatial variability of the modelled HIV prevalence across Nigeria's 37 states. In general, the variation in HIV prevalence across regions suggests that the epidemic's drivers differ in terms of population, demographic, socioeconomic, and behavioral factors.

Discussion

We built and evaluated a Bayesian predictive model that enabled HIV prevalence and burden of HIV estimation for Nigeria in March 2022. The estimated mean HIV prevalence in Nigeria was 2.1% (95% Credible Interval: 1.5–2.7%) after model diagnostics and predictive posterior model checks. Our estimate was greater than the country's national HIV prevalence of 1.4% and coincided with a previous University of Washington study's

Region	State	Spec. HIV_ Prev. 2022	NAIIS 2018 HIV Prev	HIV Prev_ (modelled) _ LCI	HIV_Prev (modelled) mid-point	HIV_Prev (modelled) _ UCI	HIV Burden_ LCI	HIV Burden midpoint	HIV Burden_ UCI	HIV Burden_ (Spectrum est. 2022)
SE	Abia	2.4	2	1.4	2	2.6	27,717	39,199	50,484	58,700
NE	Adamawa	1.2	1.1	2	2.5	3.1	39,532	50,713	62,094	33,472
SS	Akwa Ibom	4.5	4.8	2.9	3.5	4.1	83,670	100,982	118,293	173,130
SE	Anambra	2.4	2.2	2.7	3.2	3.8	81,466	99,296	116,204	96,392
NE	Bauchi	0.4	0.5	0.2	0.4	1	6120	11,949	29,144	18,524
SS	Bayelsa	1.9	1.7	2.2	2.8	3.4	28,466	36,100	43,605	24,378
NC	Benue	5.1	4.8	5	5.7	6.3	130,389	147,757	163,310	182,227
NE	Borno	1	1.1	0.8	1.4	2	22,896	41,976	58,121	38,664
SS	Cross River	2.2	1.8	2.3	2.8	3.5	45,505	56,932	69,160	52,595
SS	Delta	1.8	1.7	1.5	2.1	2.6	45,280	61,746	77,329	68,412
SE	Ebonyi	1	0.8	1.7	2.3	2.9	23,986	32,785	41,301	15,895
SS	Edo	1.9	1.8	2.9	3.4	4	64,121	76,676	89,679	39,899
SW	Ekiti	0.6	0.7	0.8	1.4	2	14,799	25,628	36,095	11,813
SE	Enugu	1.7	1.8	0.7	1.3	1.8	15,579	28,041	40,949	54,702
NC	FCT	1.2	1.4	2	2.6	3.1	54,920	69,329	85,098	45,266
NE	Gombe	1.3	1.2	0.3	0.8	1.4	4211	12,634	21,509	21,496
SE	Imo	1.8	1.7	0.4	0.9	1.5	10,400	25,860	41,882	67,383
NW	Jigawa	0.4	0.3	0.3	0.3	1	7200	8229	24,429	15,916
NW	Kaduna	1.1	1	1.7	2.3	2.8	65,790	88,639	110,700	66,417
NW	Kano	0.6	0.6	0.1	0.7	1.3	7833	43,381	80,134	49,634
NW	Katsina	0.4	0.3	0.8	1.3	1.9	26,675	46,422	65,476	21,981
NW	Kebbi	0.7	0.6	0.1	0.7	1.3	2505	14,067	25,823	20,312
NC	Kogi	1	0.8	1.5	2.1	2.7	34,201	47,526	60,185	29,066
NC	Kwara	1	0.8	1.4	2	2.6	22,185	31,247	40,777	17,149
SW	Lagos	1.2	1.3	1.2	1.8	2.4	102,008	155,155	204,874	115,346
NC	Nasarawa	1.7	1.8	2.2	2.7	3.3	25,601	32,060	38,753	61,337
NC	Niger	0.6	0.6	1.2	1.8	2.3	30,384	45,062	59,481	22,468
SW	Ogun	1.2	1.4	2.3	2.9	3.5	69,220	85,786	102,943	39,226
SW	Ondo	0.8	1	1.9	2.4	3	44,991	57,176	69,595	25,746
SW	Osun	1	0.9	1.3	1.9	2.4	31,695	45,454	59,459	31,343
SW	Oyo	0.8	0.9	0.7	1.3	1.8	15,695	28,475	40,582	43,828
NC	Plateau	1.1	1.5	1.7	2.3	2.9	35,111	46,814	57,912	32,379
SS	Rivers	3.9	3.6	4.6	5.2	5.8	188,645	213,880	240,356	195,447
NW	Sokoto	0.4	0.4	0.6	1.1	1.7	12,525	24,186	36,279	13,495
NE	Taraba	2.5	2.6	2.6	3.1	3.7	35,441	43,055	50,670	42,538
NE	Yobe	0.4	0.4	0.2	0.8	1.3	3371	12,360	20,547	7463
NW	Zamfara	0.5	0.4	0.4	0.9	1.5	7075	17,490	28,691	15,220

*Spec = Spectrum, *Prev = Prevalence, *LCI = Lower credible interval, *UCI = Upper credible interval, *ets. = Estimate. Geographical region; *NC= North Central, *NW = North West, *NE = North East, *SS = South South, *SE = South East and SW = South West.

Table 3: **Estimated HIV prevalence and HIV burden for 36 states +1 FCT in Nigeria, 26 March 2022.**

estimate of HIV prevalence for Uganda. In the University's research, Leontine et al., employed a Bayesian melding statistical technique to predict HIV prevalence, with a 95% prediction interval ranging from 2% to 7%.[32] Another study by Michael et al. used a Bayesian approach to predict HIV prevalence among gay, bisexual, and men who have sex with males in Vancouver, Canada by estimating HIV prevalence from the posterior distribution.[33] Using the estimated HIV prevalence from our Bayesian model, we determined the HIV burden in Nigeria. According to our findings, approximately 2 million people in Nigeria are living with HIV, and the HIV burden has increased by 7.2 percentage points since the last population-based survey in 2018, which is higher than the National Agency for the Control of AIDS' estimate of 1.9 million HIV-positive people.

Although the ease and flexibility with which prior information can be incorporated is a major benefit of the Bayesian approach, the development of MCMC algorithms for Bayesian computation is the primary factor responsible for the increased use and visibility of Bayesian methods in recent years.[34] One of the key advantages of the Bayesian approach is the use of posterior

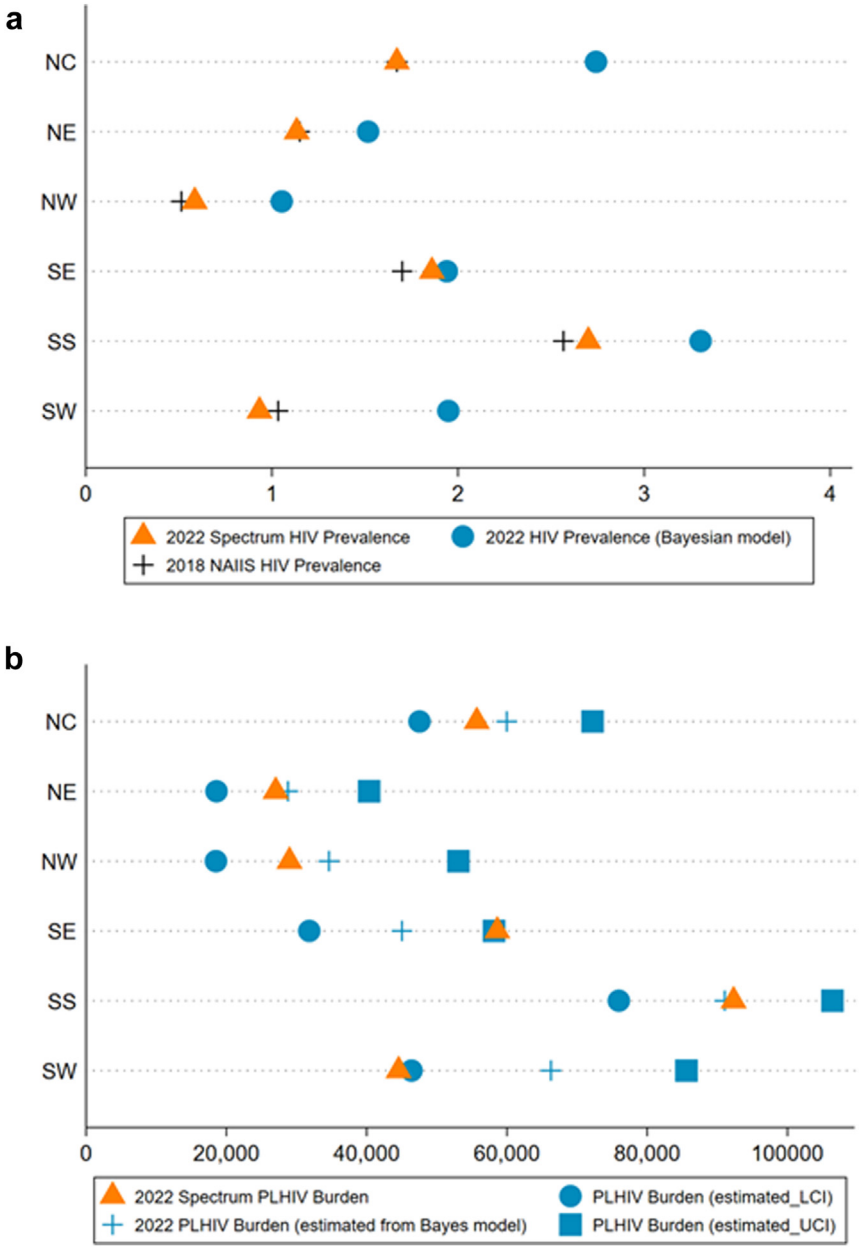

Fig. 3: **(a.)** HIV prevalence from the study (modelled prevalence) compared to the 2018 NAIIS and 2022 Spectrum estimates. The orange triangles represent the Spectrum HIV prevalence, the black cross represents the 2018 NAIIS HIV prevalence, and the blue circle represents the Bayesian model HIV prevalence. **(b.)** Overlap of the study's calculated PLHIV burden with the Spectrum projection package's estimated PLHIV burden for 2022. The PLHIV burden from the study is represented by the blue cross, while the HIV burden for the lower and upper credible intervals is represented by the blue circle and square, respectively. **Geographical region**: *NC= North Central. *NW = North West. *NE = North East. *SS = South South. *SE = South East. SW = South West. **Credible intervals:** *LCI = Lower credible interval. *UCI = Upper credible interval. Overall, the modelled HIV prevalence was close to the 2018 NAIIS HIV prevalence, but the large differences in HIV prevalence between the 2018 NAIIS and modelled estimates, particularly in the NC and SW, indicate differences in the target population tested for HIV. The PLHIV estimates derived from the spectrum appear to be well contained within the lower and upper credible intervals estimated by the Bayesian. This means that the estimated PLHIV burden for Nigeria overlaps with the PLHIV estimates predicted by Spectrum software, highlighting the accuracy of the Bayesian model estimates.

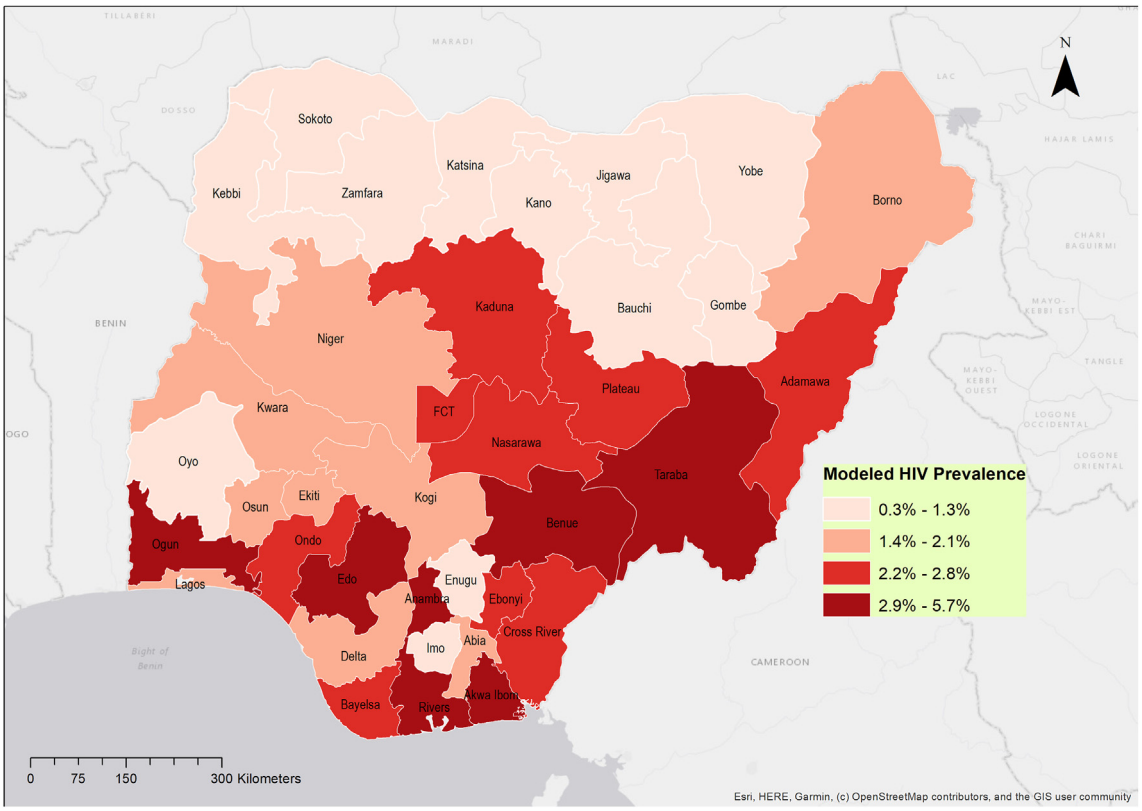

Fig. 4: Choropleth map displaying HIV Prevalence spread across Nigeria.

probability. Posterior means and 95% credible intervals were used to summarise our model parameters (Supplementary Material 2). The effective sample size drawn from the posterior distribution in our model was 4000, which was extremely close to the nominal sample size. The autocorrelation at 0 was 1, and it quickly fell off after that (Supplementary Material 3). There was no warning on diverging transitions, and according to the potential scale reduction statistic or the Gelman-Rubin convergence diagnostic known as "*Rhat*", all model parameters were less than 1.1, and this indicates that all chains converged to the same stationary target distributions at the same time.

In Nigeria, the first HIV Sentinel Survey in 1991 showed a prevalence of 1.8%. Subsequent sentinel surveys produced prevalence of 3.8% in 1993, 4.5% in 1996, 5.4% in 1999, 5.8% in 2001, 5.0% in 2003, 4.4% in 2005, 4.6% in 2008, 4.1% in 2010 and 1.4% in 2018 (Supplementary Material 1). This statistic shows that HIV prevalence in Nigeria has steadily declined between 1999 and 2018. Our findings reveal that new HIV infections are on the rise, which could be attributed to key populations, their clients, and sexual partners, who accounted for 64% of all new HIV infections in West and Central Africa and 25% in the East and Southern African subregion.[35]

The subnational heterogeneity of HIV prevalence has been demonstrated by previous studies in Uganda,[36] South Africa,[37] and Nigeria.[38] In *The Lancet HIV*, Larry Chang and colleagues[39] use a population-based cohort study to analyze HIV disease burden, sexual behaviors, and treatment and prevention service coverage in Rakai, Uganda, and documented significant heterogeneities. In agrarian (n = 9931), trading (n = 3318), and fishing (n = 3870) communities, they mapped HIV prevalence and examined differences in HIV risk factors, antiretroviral medication uptake, and male circumcision. HIV prevalence ranged from 9% to 43%, with Lake Victoria fishing settlements having the greatest prevalence. Data examined by Larry Chang et al., were similar to the input data on HIV risk factors, treatment and prevention data included in our Bayesian model to estimate HIV prevalence. As of 2008, NAIIS survey data provided an overview of the national HIV epidemic. In our study, we used data from the national HIV program reported in 2022. Over the last 4 years, prevention and treatment services have been intensified in targeted regions across the country, which the authors believe can explain the predicted decline in HIV prevalence. It was found in a recent longitudinal study by Joseph Kagaayi[40] et al., published in the Lancet that after five years of increasing HIV testing coverage, male circumcision coverage, ART

coverage, and population HIV viral load suppression, HIV incidence decreased from 3.43/100 person-years to 1.59/100 person-years. Moreover, a previous Lancet study found that the model projections accurately predicted the 1.6 percentage point decline in prevalence among 15-24-year-olds when compared to household survey data. To quantify the impact of interventions on the HIV epidemic, the authors recommend that similar analyses be conducted in Nigeria. In our study, HIV prevalence was highest in the South South (3.30%, 95% CI: 2.7–3.9) compared to other regions, followed by North central (2.74%, 95% CI: 2.14–3.31) and the prevalence was similar in the South West (1.95%, 95% CI: 1.37–2.52) and South East (1.94%, 1.38–2.52), and lowest in the North West (1.0%, 95% CI: 0.57–1.64). Our findings collaborate the results of the 2018 NAIIS report, which found that the South–South zone of the country had the highest HIV prevalence, with 3.1% of persons aged 15–49 years infected. The North Central zone (2.0%) and the South East zone (1.9%) also have high HIV prevalence. The South West zone (1.1%), the North East zone (1.1%), and the North West zone (0.6%) have lowest HIV prevalence.[41] The South–South zone represents state in the Niger delta region (i.e., a densely populated and petroleum-rich region that has become the focus of international concern over pollution caused primarily by catastrophic oil spills by multinational businesses in the petroleum industry[42]). The high HIV prevalence in the Niger Delta of Nigeria is generally attributed to concurrent sexual partnerships, weak public sector health care and education systems, poverty, migration, and sex work.[43]

Our study suggests an inverse relationship between poverty (wealth index for the poorest), fertility and literacy rate for women with HIV. According to the Bayesian regression analysis, wealth index for the wealthiest people rises when HIV prevalence rises, while wealth index for the poorest people rises as HIV prevalence decreases. Contrary to popular belief that poverty acts as an underlying driver of HIV infection in sub-Saharan Africa, an increasing body of research at the national and individual levels reveals that wealthier countries, as well as wealthier individuals within countries, are at higher risk for HIV infection.[44]

Recent studies report that longer HIV infection is associated with greater relative fertility reduction for HIV-positive women.[45] In a meta-analysis study by James et al., a mathematical model was used to demonstrate the impact of the HIV/AIDS epidemic on the number of births in Uganda. Fertility was lower among HIV-infected women than HIV-uninfected women, except for those aged 15–19 years, in whom the selective pressure of sexual debut on pregnancy and HIV infection led to higher fertility rates among the HIV infected. According to our findings, as fertility rises in HIV-uninfected women, fertility falls by 0.54 in HIV-positive women. Our findings show that HIV has a population-level effect on fertility, which has crucial implications for future assessments of the requirement for a national PMTCT program.

In terms of literacy rate, our research shows that as women's literacy increases, HIV prevalence reduces by 0.06. Our findings support Julia Andrea's research, which used the 2010 Malawi Demographic Health Survey and the 2011 Uganda AIDS Indicator Survey in a regression discontinuity model to investigate the causal relationship between primary schooling and adult HIV status in Malawi and Uganda, two countries in South-eastern and Eastern Africa respectively with some of the world's highest HIV infection rates. In Malawi, a one-year increase in schooling reduces the likelihood of an adult woman testing positive for HIV by 0.06 ($p < 0.01$) and in Uganda, by 0.03 ($p < 0.05$).[46]

This research has several limitations. The Bayesian normal linear regression analysis was used to model the relationship between variables, with the assumption that the observations follow a normal distribution. Literature indicates that GLM can be used as an alternative method for modeling HIV prevalence estimates.[47] Modeling proportional data directly is possible with the GLM framework based on binomial likelihood and logit-link. Future analysis to improve the predicted HIV prevalence estimates at state level will use the GLM framework with binomial likelihood and logit-link. Fig. 2 in Supplementary Material suggests that there are minor deviations between the observed and predicted data, which could be due to outliers and may have influenced model results. One of the major biases that can occur when using program data to estimate population level HIV prevalence is the problem with non-random distribution of the HIV positive cases identified on the program. This is because testing for HIV on the national HIV program is driven by epidemiology in terms of areas of high HIV prevalence, high HIV burden, or prevalence of risk behaviors that increases the chances of HIV acquisition and transmission. HIV testing is also encouraged through community outreach programs run by healthcare professionals or HIV counselors and testers, as well as by testing in designated hotspots where key populations gather. In population prevalence surveys, HIV testing participants are chosen at random from a subset of the community using a random sample technique. In addition, in Supplementary Material 2, the variable on the number of pregnant women who tested HIV negative indicated a negative prediction value. While the negative predicted values were limited to the minimum values, the other five-number summary descriptive parameters (1st quartile, median, mean, 3rd quartile and maximum) indicated positive values. The authors presented the 1st quartile (as lower credible interval), median (point estimate) and 3rd quartile (as upper credible interval) in our study. The results of the model may also have been influenced by this.

Based on the Nigerian HIV program and data used in this study, the Bayesian model's HIV prevalence estimates might also have been impacted by non-random selection of individuals for HIV testing. As part of this study, we included data from both the facility level and the community level regarding HIV testing. The facility provided HIV testing data from service delivery points (SDPs). These SDPs included testing in a prevention of mother-to-child transmission unit, a tuberculosis unit, an emergency ward, an inpatient and surgery ward, an STI clinic, and contact tracing and testing of index cases. Among the methods of testing used in the community were mobile testing, voluntary counseling, and testing (VCT), and index case testing. The testing strategy was largely targeted at the general population and priority populations such as: pregnant women, key populations, adolescent girls, and young women and may have contributed to the upward bias in prevalence and key differences between our modelled PLHIV estimate and that from NAIIS.

In Nigeria, antiretroviral therapy coverage (the percentage of people living with HIV receiving antiretroviral therapy) at the state level is used to evaluate the progress of the national HIV/AIDS program implementation and the gaps in care across regions and populations (adults, pediatrics, and KPs). ART coverage among adults has exceeded 100% in several regions. Due to the limited quality of the population dataset used in the country's annual HIV projections, this anomaly arises. National PLHIV estimations are based on the 2006 Nigeria population census, which had a population of approximately 140 million. UN data shows that Nigeria's population is expected to reach 218 million in 2022, an increase of 55% over the last 16 years. Although the country's model uses the 2006 population dataset for its annual projection, it is evident that this introduces bias in the total number of PLHIV estimated since different regions in Nigeria continue to identify and treat more PLHIVs than were originally estimated, resulting in more than 100% ART coverage. Our model incorporated ART coverage data, which might have influenced the overall model results. In the national program, PMTCT coverage data are used to track progress toward ensuring that pregnant women who attend antenatal care (ANC) know their HIV status and are initiated on ART. During one pregnancy, a woman could be tested multiple times; therefore, the national program ensures that a sound data collection and reporting system is in place to minimize double counting, including a longitudinal ANC registry. In one of the states, PMTCT coverage was greater than 100%, which may have affected our analysis.

Additionally, in this study, the authors presume that the covariates derived from the 2017 MICS and 2018 NDHS are relatively stable over time.

The Bayesian statistical model employed the MCMC approach to randomly create posterior estimates, which were then used to forecast the new HIV prevalence at state level. Missing HIV program data in Anambra and Ebonyi states that were not covered by the national PEPFAR program were imputed using the predictive mean matching (PMM) multiple imputation technique. Multiple imputation, on the other hand, incorporates all available data, preserves sample size and statistical power, and produces unbiased estimates with greater validity than ad hoc techniques to missing data. We only had quantitative data and did not collect qualitative data that would have helped explain the geographical heterogeneity of HIV prevalence across the 36 states +1 FCT. Census data used was from U.S. Census Bureau subnational population projections for 2020. Population census estimates sourced from US Census Bureau database gave more reliable and accurate population figures unlike using the 2.6% annual growth rate that is usually applied to the 2006 Nigeria census data which may lead to bias. Given the COVID-19 pandemic, HIV testing services data collected from October 1, 2020 to September 30, 2021, may not be representative. However, according to the national HTS program data, almost 10 million people (adults and children) were tested for HIV during this time, representing 5% of the whole Nigerian population as of 2020, compared to the total sample size of 172,000 utilised in the 2018 NAIIS, which was less than 1% of the total Nigerian population.

Our Bayesian predictive model is a valuable addition to the toolkit for estimating HIV prevalence using national program data and surveys. This model provides more comprehensive and flexible use of evidence to estimate state-level HIV prevalence for Nigeria using program data and adjusting for explanatory variables. This methodology can be used to evaluate potential problems with conventional HIV prevalence estimates. The use of a Bayesian statistical approach to combine multiple datasets into a quantitative framework has a lot of potential for HIV epidemiology model fitting, and it may be utilised to enhance national integrated surveillance initiatives for infectious diseases response in Nigeria.

Contributors

AAO conceived and designed the study and developed the methodology, principal in data management, did the data analysis, and wrote the manuscript draft. AA provided technical oversight and validation of the data collection process and methods utilised. DO performed critical reviews of the first and final versions of the manuscript. MK, BK, MD, and LL provided peer-review support and reviewed codes used to run the Bayesian statistical model in R-software. AAO, AA, BK and MK verified the datasets. DP provided technical reviews and proofreading of the manuscript. When the study's findings were presented to the Nigerian government, BA, GA, and OO contributed to reviews. RG and HM provided project administration oversight and validation of the program data. All the authors; AAO, AA, DO, MK, BK, MD, LL, DP, BA, GA, OO, RG, and HM accept responsibility for the decision to submit for publication".

Data sharing statement

These aggregate level prevention and treatment program data from the national HIV program aren't publicly accessible but may be obtained upon request. The 2018 NAIIS report, 2018 NDHS and 2016/2017

MICS data for Nigeria is publicly accessible. The code required to replicate the Bayesian statistical model, test and analysis would be available upon request to the authors.

Declaration of interests

In this study, we report no financial or non-financial competing interests. All authors report no disclosures. The contents in this study are those of the authors and do not necessarily reflect the view of the U.S. President's Emergency Plan for AIDS Relief, the U.S. Agency for International Development or the U.S. Government.

Acknowledgements

The authors would like to express their gratitude to those who participated at meetings of the US Agency's Strategic Information Technical Working Group in Nigeria and the UNAIDS Reference Group on Estimates, Modelling, and Projections, where the Bayesian statistical model's findings were presented. There feedback was very relevant and added value to our study. This article was made possible by the support of the American people through the United States Agency for International Development (USAID) under the U.S. President's Emergency Plan for AIDS Relief (PEPFAR).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2023.102098.

References

1 Global aids response–NACA Nigeria (n.d.). Retrieved March 30, 2022, from: https://naca.gov.ng/wp-content/uploads/2016/11/Nigeria_GARPR_2015_Report_0.pdf.
2 Sheehy M, Tun W, Vu L, et al. High levels of bisexual behavior and factors associated with bisexual behavior among men having sex with men (MSM) in Nigeria. *AIDS Care*. 2014;26:116. from: https://pubmed.ncbi.nlm.nih.gov/23742663/. Accessed March 30, 2022.
3 Djomand G, Quaye S, Sullivan PS. HIV epidemic among key populations in West Africa. *Curr Opin HIV AIDS*. 2014;9:506. from: https://pubmed.ncbi.nlm.nih.gov/25010898/. Accessed March 30, 2022.
4 Global report 2013–UNAIDS (n.d.). Retrieved March 30, 2022, from: https://files.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf.
5 Understanding fasttrack–UNAIDS (n.d.). Retrieved March 12, 2022, from: https://www.unaids.org/sites/default/files/media_asset/201506_JC2743_Understanding_FastTrack_en.pdf.
6 UNAIDS. *Global report: UNAIDS report on the global AIDS epidemic 2010*. Geneva: UNAIDS; 2010.
7 Brown T, Salomon JA, Alkema L, et al. Progress and challenges in modelling country-level HIV/AIDS epidemics: the UNAIDS Estimation and Projection Package 2007. *Sex Transm Infect*. 2008; 84(Suppl 1):i5–i10.
8 Schwartlander B, Stover J, Walker N, et al. AIDS. Resource needs for HIV/AIDS. *Science*. 2001;292:2434–2436.
9 Salomon JA, Hogan DR, Stover J, et al. Integrating HIV prevention and treatment: from slogans to impact. *PLoS Med*. 2005;2:e16.
10 Shana O, Rehle T, Simbayi LC, et al. *South African national HIV prevalence, incidence, behaviour, and communication survey 2008: a turning tide among teenagers?* Cape Town: Human Sciences Research Council; 2009.
11 Central Statistical Office (CSO), Ministry of Health (MOH), Tropical Diseases Research Centre (TDRC), et al. *Zambia demographic and health survey 2007*. Calverton, Maryland, USA: CSO and Macro International Inc; 2009.
12 National AIDS Coordinating Agency (NACA), Central Statistics Office (CSO) and Other Development Partners. *The Botswana AIDS impact survey II (BAIS II): popular report*. Gaborone: National AIDS Coordinating Agency; 2005.
13 Boerma JT, Ghys PD, Walker N. Estimates of HIV-1 prevalence from national population-based surveys as a new gold standard. *Lancet*. 2003;362:1929–1931.
14 Ghys PD, Walker N, McFarland W, et al. Improved data, methods and tools for the 2007 HIV and AIDS estimates and projections. *Sex Transm Infect*. 2008;84(Suppl 1):i1–i4.
15 Bärnighausen T, Bor J, Wandira-Kazibwe S, et al. Correcting HIV prevalence estimates for survey nonparticipation using Heckman-type selection models. *Epidemiology*. 2011;22:27–35.
16 Mishra V, Barrere B, Hong R, et al. Evaluation of bias in HIV seroprevalence estimates from national household surveys. *Sex Transm Infect*. 2008;84(Suppl 1):i63–i70.
17 Nigeria indicator and technical impact survey report–CIHEB (n.d.). Retrieved March 12, 2022, from: http://ciheb.org/media/SOM/Microsites/CIHEB/documents/NAIIS-Report-2018.pdf.
18 Weiser SD, Heisler M, Leiter K, et al. Routine HIV testing in Botswana: a population-based study on attitudes, practices, and human rights concerns. *PLoS Med*. 2006;3:e261.
19 Kalichman SC, Simbayi LC. HIV testing attitudes, AIDS stigma, and voluntary HIV counselling and testing in a black township in Cape Town, South Africa. *Sex Transm Infect*. 2003;79:442–447.
20 Kranzer K, McGrath N, Saul J, et al. Individual, household and community factors associated with HIV test refusal in rural Malawi. *Trop Med Int Health*. 2008;13:1341–1350.
21 Reniers G, Eaton J. Refusal bias in HIV prevalence estimates from nationally representative seroprevalence surveys. *AIDS*. 2009; 23:621–629.
22 Session VII: National Population Projections–Un (n.d.). Retrieved March 12, 2022, from: https://www.un.org/en/development/desa/population/events/pdf/other/14/7-BuildingPopulationProjection.pdf.
23 *Nigeria population (live)*. Worldometer; 2023. from: https://www.worldometers.info/world-population/nigeria-population/. Accessed March 13, 2022.
24 Bureau, U. S. C.. *Subnational population by sex, age, and geographic area*. Census.gov; 2022. from: https://www.census.gov/geographies/mapping-files/time-series/demo/international-programs/subnationalpopulation.html. Accessed March 13, 2022.
25 (HTS_Index) number of individuals who were identified and tested using index testing services and received their results. (HTS_Index) Number of individuals who were identified and tested using Index testing services and received their results | Indicator Registry (n.d.). Retrieved October 17, 2021, from: https://indicatorregistry.unaids.org/indicator/htsindex-number-individuals-who-were-identified-and-tested-using-index-testing-services.
26 Mason K, , Johns Hopkins Bloomberg School of Public Health, Ketende S, et al. *A cross-sectional analysis of population demographics, HIV knowledge and risk behaviors, and prevalence and associations of HIV among men who have sex with men in the Gambia*. Mary Ann Liebert, Inc.; 2013. https://www.liebertpub.com/doi/10.1089/aid.2013.0092.
27 Vuylsteke B, Semde G, Sika L, et al. *High prevalence of HIV and sexually transmitted infections among male sex workers in Abidjan, côte d'ivoire: need for services tailored to their needs*. Sexually Transmitted Infections; 2012. https://sti.bmj.com/content/88/4/288.
28 Coftadeh. United Nations Office on Drugs and Crime. HIV and AIDS (n.d.) Retrieved August 31, 2022, from: https://www.unodc.org/nigeria/en/hiv-and-aids.html#:~:text=The%202019%20Nigeria%20National%20HIV,who%20inject%20drugs%20(PWIDs).
29 Zaba B, Gregson S. Measuring the impact of HIV on fertility in Africa. *AIDS*. 1998;12(Suppl 1):S41–S50. PMID: 9677188.
30 *Probabilistic projections of HIV prevalence using Bayesian melding*. JSTOR; 2007. from: https://www.jstor.org/stable/4537430. Accessed March 28, 2022.
31 Irvine MA, Konrad BP, Michelow W, Balshaw R, Gilbert M, Coombs D. A novel Bayesian approach to predicting reductions in HIV incidence following increased testing interventions among gay, bisexual and other men who have sex with men in Vancouver, Canada. *J R Soc Interface*. 2018;15:20170849. from: https://royalsocietypublishing.org/doi/10.1098/rsif.2017.0849. Accessed March 29, 2022.
32 Craig BA, Fryback DG, Klein R, et al. A Bayesian approach to modeling the natural history of a chronic condition from observations with intervention. *Stat Med*. 1999;18:1355–1371.
33 Arminger G, Muthen BO. A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika*. 1998;63:271–300.
34 *The global HIV/AIDS epidemic*. KFF; 2021. from: https://www.kff.org/global-health-policy/fact-sheet/the-global-hivaids-epidemic/#footnote-513952-8. Accessed March 29, 2022.
35 Chang LW, Grabowski MK, Ssekubugu R, et al. Heterogeneity of the HIV epidemic in agrarian, trading, and fishing communities in Rakai, Uganda: an observational epidemiological study. *Lancet HIV*.

2016;3:e388–e396. from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4973864/. Accessed March 29, 2022.
36 Tanser F, Lesueur D, Solarsh G, Wilkinson D. HIV heterogeneity and proximity of Homestead to roads in rural South Africa: an exploration using a geographical information system. *Trop Med Int Health*. 2000;5:40. from: https://pubmed.ncbi.nlm.nih.gov/10672204/. Accessed March 29, 2022.
37 Odimegwu CO, Alabi O, De Wet N, Akinyemi JO. Ethnic heterogeneity in the determinants of HIV/AIDS stigma and discrimination among Nigeria women–BMC public health. *BMC Public Health*; 2018. from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-018-5668-2. Accessed March 29, 2022.
38 Define_me (n.d.). Retrieved March 29, 2022, from: https://www.thelancet.com/journals/lanhiv/article/PIIS2352-3018(16)30048-0/fulltext.
39 An appraisal of issues in the Niger Delta (n.d.). Retrieved March 30, 2022, from: https://globalacademicgroup.com/journals/academic%20scholarship/ANA%20APPRAISAL%20OF%20ISSUES%20IN%20THE%20NIGER%20DELTA.pdf.
40 Ratmann O, Kagaayi J, Hall M, et al. Quantifying HIV transmission flow between high-prevalence hotspots and surrounding communities: a population-based study in Rakai, Uganda. *Lancet HIV*. 2020;7:e173. https://pubmed.ncbi.nlm.nih.gov/31953184/.
41 Udoh IA, Mantell JE, Sandfort T, Eighmy MA. *Potential pathways to HIV/AIDS transmission in the Niger Delta of Nigeria: poverty, migration and commercial sex*. AIDS Care; 2009. from: https://www.ncbi.nlm.nih.gov/sites/ppmc/articles/PMC2879970/. Accessed March 29, 2022.
42 Shelton J, Cassell M, Adetunji J. Is poverty or wealth at the root of HIV? Semantic scholar. undefined. *Lancet*; 1970. from https://www.semanticscholar.org/paper/Is-poverty-or-wealth-at-the-root-of-HIV-Shelton-Cassell/702448c722bce789c0b374d8033df827740168b7. Accessed March 29, 2022.
43 Ogunmola OJ, Oladosu YO, Olamoyegun MA. *Relationship between socioeconomic status and HIV infection in a rural: HIV*. HIV/AIDS–Research and Palliative Care; 2014. from: https://www.dovepress.com/relationship-between-socioeconomic-status-and-hiv-infection-in-a-rural-peer-reviewed-fulltext-article-HIV. Accessed March 29, 2022.
44 *(PDF) A study of the association of HIV infection with wealth in sub-Saharan Africa*; 2007. from: https://www.researchgate.net/publication/237534201_A_study_of_the_association_of_HIV_infection_with_wealth_in_sub-Saharan_Africa. Accessed March 29, 2022.
45 Lewis JJ, Ronsmans C, Ezeh A, Gregson S. The population impact of HIV on fertility in sub-Saharan africa. *AIDS (London, England)*. 2004;18:S35. from: https://pubmed.ncbi.nlm.nih.gov/15319742/. Accessed March 29, 2022.
46 Ganguly S, Goswami DN, Mondal S, Chakrabarti S, Mundle M. A retrospective cohort study on effect of literacy status of HIV-positive pregnant women on possibility of child getting HIV infected. *J Fam Med Prim Care*. 2018;7:167. from: https://www.ncbi.nlm.nih.gov/pubmed/29915753. Accessed March 29, 2022.
47 GLMS: Binomial Data. Introduction to R (n.d.). https://www.simonqueenborough.info/R/statistics/glm-binomial