Modelling Spatial Variations of HIV Prevalence in Nigeria: Implications for Effective Prevention Strategies

Samson B. Adebayo*
National Agency for Food & Drug Administration and Control, Wuse Zone 7, Abuja, Nigeria
samson.adebayo@nafdac.gov.ng

Ezra Gayawan
Department of Statistics, Federal University of Technology, Akure, Nigeria
egayawan@futa.edu.ng

Adeniyi F. Fagbamigbe
Department of Epidemiology and Medical Statistics, College of Medicine, University of Ibadan, Nigeria
franstel74@yahoo.com

Fatai W. Bello
Department of Health Planning & Management
University of Ilorin, Ilorin, Nigeria
fwbello@gmail.com

Abstract

Understanding the correlates of HIV epidemics and its risk factors plays a vital role in designing appropriate strategies to mitigate its scourge. Availability of reliable data on the spread of HIV in the general population was until recently a challenge. Hence estimates were based on data from HIV surveillance and sentinel studies among pregnant women attending antenatal clinics at some selected sentinel sites. Such data were found not to be a true representation at population-based especially in countries with generalized epidemics.

This study explored possible association of respondents' marital status and geographical variations with HIV prevalence among the general population of male and female based on two waves of the National HIV/AIDS and Reproductive Health Survey conducted in Nigeria in 2007 and 2012. Correlates and possible risk factors of HIV are also explored to provide guidance for programmers, policy makers and donors of HIV and other sexual reproductive health.

Bayesian Geoadditive modelling technique based on structured additive approach was employed for analysis. Findings reveal pronounced and significant spatial variations at a highly disaggregated level of states in Nigeria. When these findings are taken into consideration in designing intervention strategies, it is believed that each state can be targeted with the right intervention(s). This in turn will results in efficient utilization of the scarce resources witnessed globally and more importantly with the economic recession in Nigeria.

Keywords: reliable data; appropriate strategy; Markov Chain Monte Carlo techniques; Nigeria.

1. Introduction

A proper understanding of the spatial component of Human Immunodeficiency Virus (HIV) among the administrative units in a country is key to structuring, developing and implementing apposite strategies that will have an effect on people at the local administrative units (Manda, Lombard, Mosala, 2012).

Reliable data on the spread of HIV and its risk factors in the general population are essential for an effective response to the epidemic and its consequences. In countries with generalized epidemics, national estimates of HIV prevalence and trends in the adult population are generally derived indirectly from HIV surveillance among pregnant women attending selected antenatal clinics. (WHO & JUNPA, 2003; Saphonn *et al*., 2002; Zaba *et al*., 2000; JUNPA & WHO, 2005). However, of recent, population based surveys that allows for studying the prevalence and other risk factors at more localized levels have emerged in many of these countries.

An understanding of culture among different geographical locations is crucial to a study of HIV and the global variations in its prevalence. It is frequently assumed that affluent countries have escaped the strongest influence of the virus because they have both health care and educational systems that prevent the spread of epidemics. However a look at some of the basic history of the HIV shows that this is not the case. The first crisis of HIV and the disease it causes, Acquired Immune Deficiency Syndrome (AIDS) occurred in the USA in the 1980s, one of the richest countries in the world (MacQueen, 1994).

Today Sub-Saharan Africa is in the midst of a HIV/AIDS crisis, and with it the stigma that AIDS is a virus of poverty. HIV prevalence in Nigeria, estimated from the population-based study varies significantly by sex, marital status, age and sexual risk behaviours. A general population survey conducted in 2007 put national HIV prevalence at 3.6%, and the prevalence was higher among females (4.0%) than males (3.2%) (FMOH, 2008). The latest of the national surveys in 2012 showed a slightly lower general population HIV prevalence of 3.4% (95% CI: 3.2%-3.6%) (FMOH, 2013). Findings from Adebayo *et al*. (2013) and Fagbamigbe, Adebayo & Idemudia (2016) showed that HIV prevalence is about twice among women of reproductive age compared with their male counterparts. Further, HIV prevalence among women who were formerly married (i.e. divorced, widowed and separated) is about three times their counterparts who were currently married or cohabiting with sexual partners. Also, HIV prevalence among men and women who engaged in a transactional sex is more than twice as likely as those who had never engaged in a transactional sex.

Similar to other studies, assumption of linear effect for metrical covariates such as age has been shown to be too rigid. HIV prevalence by age has a nonlinear relationship, assuming an inverted U-shape (Jonhnson & Way, 2006; Mishra, Montana and Neuman, 2007). In particular, HIV prevalence is low among people below the age of 18, and increases up to the age of 35–40 then starts declining with increase in age (Ngesa, Mwambi and Achia, 2014). But studies that demystify the spatial pattern particularly after accounting for the influence of relevant risk factors have not been adequately undertaken. Such studies would provide information hitherto unravelled by survey reports, and that could be useful for effective intervention strategies.

In this study, we employed a flexible geoadditive modelling of HIV prevalence among the general population based on the two waves of the National HIV/AIDS Reproductive Health Survey conducted in 2007 and 2012. Bayesian approach based on Markov Chain Monte Carlo (MCMC) techniques was adopted. Appropriate priors are assigned on relevant parameters.

2. Survey Methodology and Data

Data from the 2007 and 2012 National HIV/AIDS and Reproductive Health and Serological Surveys (NARHS Plus) were used. These were cross-sectional surveys of men (15 – 64 years) and women (15 – 49 years) of reproductive age living in households in rural and urban areas in all 36 states and the Federal Capital Territory (FCT) of Nigeria. The survey contained both behavioural and serological components. Multi-stage cluster sampling techniques were used to select eligible persons. Further sampling and data collection details can be found in FMOH, 2013. Globally approved standard procedures for HIV counselling, testing and result validation were adopted in the survey as reported in FMOH, 2013.

2.1. Data

This paper considered women respondents who participated in the behavioural component of the survey and consented to HIV testing. All analyses involving HIV status in the present study were based on the 4,195 (2007) and 11,946 (2012) women with both behavioural survey results and valid HIV test results.

Dependent variable: The outcome variable of interest in this study was the HIV sero-status test result obtained from female respondents who were interviewed and consented to HIV testing. A reactive HIV test was coded ‘1’ and non-reactive outcomes as ‘0’.

Independent variable: These include marital status, rural-urban location of residence, state (geographical locations) of residence, educational attainment, geopolitical zones, age, self-reported sexual behaviour within the l2 months preceding the survey such as multiple partnering, transactional sex, age at first sex, experience of sexually transmitted infections (STIs) and comprehensive knowledge of HIV as used in earlier studies. Marital status was grouped into 'never married', 'currently married/cohabiting', and 'formerly married' (i.e., separated, divorced, or widowed).

3. Data Analysis

3.1 Model

Suppose the observations $(y_{ijk}, z_k, v_k, r_j, s_i)$, where y_{ijk} is the HIV sero-status of k^{th} woman living in the j^{th} community in i^{th} district/state, vector $z=(z_{k1}, \ldots, z_{kp})'$ of metrical covariates (say age of the respondent in years), vector $v = (v_{k1}, \ldots, v_{kq})'$ of categorical covariates and $s_i = (1,\ldots, 37)$ are the community and state (district) where the k respondent lived during the survey. The distribution of the response variable y belongs to the exponential family with mean $\mu = E(y|z,\beta)$ which can be linked to a linear predictor η thus:

$$\mu = h(\eta) \quad \text{where} \quad \eta = z'\beta$$

h is a known response function and β unknown regression parameters. Our intention is to jointly model the dependence of y_{ijk} on metrical, spatial and categorical covariates within the context of generalized additive model (Hastie and Tibshirani, 1990). The predictor η_{ijk} for the geo-additive model is defined as

$$\eta_{ijk} = time_2 + \sum_{j=1}^{p} f_{ijk}(z_{ijk}) + f_{spat}(s_i) + v_{ijk}\beta \qquad (1)$$

where f is the nonlinear unknown smooth function of the metrical covariates, f_{spat} is the nonlinear spatial effect, β is the regression coefficients for the categorical covariates and $time_2$ is the effect of second round of the NARHS data: 2012 (i.e. year of study with 2007 as the reference category).

3.2 Bayesian Inference

A full Bayesian approach was adopted for estimation of parameters. In Bayesian analysis, the proposed model of the observed data is combined with the prior distribution of all the unknown model parameters. Independent diffuse priors are assumed on the parameters of fixed effects though a weakly informative normal prior would also have been possible. For the non-linear effects, a Bayesian P-splines prior based on Lang and Brezger (2004) and Brezger and Lang (2006) was assumed. The P-splines allows for nonparametric estimation of f as a linear combination of basis function (B-splines)

$$p(z) = \sum_{t=1}^{m} \alpha_t B_t(z)$$

where $B_t(z)$ are B-splines and the coefficients α_t are further defined to follow a first or second order Gaussian random walk smoothness priors

$$\alpha_1 = \alpha_{t-1} + \varepsilon_1 \qquad \alpha_1 = 2\alpha_{j-1} - \beta_{j-2} + \varepsilon_1$$

with *i.i.d.* errors $\varepsilon_t \sim N(0, \tau^2)$. The variance τ^2 controls the smoothness of f. Assigning a weakly informative inverse Gamma prior $\tau^2 \sim IG(\epsilon, \epsilon)$, ϵ small, it is estimated jointly with the basis function coefficients. For the geographical effects $f_{spat}(s)$, $s=1, \ldots, S$, we assume a Gaussian Markov random field prior. Basically, this is an extension of first order random walk priors to two-dimensional spatial arrays, see Rue and Held (2005) for general information.

For the spatial effects $f_{str}(s)$, we chose a Gaussian Markov random field prior (2) which is common in spatial statistics, see Besag *et al*. (1991).

$$(f_{str}(s) \mid f_{str}(t); t \neq s, \tau^2) \sim N\left(\sum_{t \in \partial_s} f_{str}(t)/N_s, \tau^2/N_s\right) \qquad (2)$$

In order to estimate the smoothing parameters for non-linear and spatial effects simultaneously, highly dispersed but proper hyper-priors were assigned to them. Hence for all variance components, an inverse gamma distribution with hyperparameters a and b is chosen, e.g. $\tau^2 \sim \mathrm{IG}(a,b)$. Standard choices of hyperparameters are a=1 and b=0.005 or a=b=0.001. Basis of inference in this study is carried out through computationally efficient Markov chain Monte Carlo (MCMC) techniques.

4. Results and Discussions

For the three models considered, findings show varying degree of significance based on the linear effects. While a decline in HIV prevalence between 2007 and 2012 was not significant in the model with unsplit spatial effect, it was significant in the models without spatial effect or with spatial effect being split into structured and unstructured effect. A somewhat positive association was evident in educational attainment in models (b) and (c), showing higher odds of contracting HIV for respondents who attained at least a primary educational level. Wealth quintiles do not show any association with HIV prevalence. Respondents who were formerly married are about two times more likely to be HIV positive compared with those who were never married or still single as at the time of surveys.

Figure 1 (a & b) shows findings on nonlinear effects of geographical locations. The left panel shows the map of estimated total spatial effects (i.e. structured and unstructured) while the centre panel presents the map of significance. States with black colour in the centre panel indicate states with negative credible intervals i.e. states that are significantly associated with lower prevalence. Similarly, states with white colour imply states with positive credible intervals i.e. states that are significantly associated with higher HIV prevalence. The results show that Benue, Nasarawa, FCT, Rivers and Akwa Ibom states are locations where the likelihood of contracting HIV are significantly higher but lower in Kebbi and Zamfara states. The estimates are not significant in the other states shaded in grey colour. The right panel presents the findings of the nonlinear effect of respondents' age. An almost inverse 'V' shape was evident with respondents in age group 25 – 35 years being mostly affected with the scourge of the virus. This is similar to findings from other studies on HIV prevalence in a population-based study.

5. Conclusions

Evidently, a strong association between respondents' marital status, educational attainment and HIV prevalence exists. Findings on age demonstrate that the risk of HIV is peak among women in the neighbourhood of 30 years of age. This is evidently the age where fertility and hence sexual activities is considered peak in most African countries (Gayawan et al., 2010) and with would therefore be worthwhile if these women are targeted in any intervention programme. The prevalence of HIV also varies considerably by the respondents' geographical locations. This paper has provided insight into an effective utilization of the scarce resources which can be achieved through cost-effective HIV prevention strategies by proper prioritization of needs.

References

Adebayo, S.B., Olukolade, R.B., Idogho, O., Anyanti, J. & Ankomah, A. (2013). Marital Status and HIV Prevalence in Nigeria: Implications for Effective Prevention Programmes for Women. *Advances in Infectious Diseases,* 3: 210-218.

Adebayo, S.B., Anyanti, J., Ladipo, O., Ankomah, A. & Ekweremadu, B. (2009) Understanding the relationship between HIV/AIDS Gender and Marital Status: Evidence from national population based HIV testing in Nigeria. *American Public Health Association conference, Philadelphia, U.S.A. (4th November, 09 to 11th November, 09).*

Besag, J., York, Y., Mollie, A. (1991). Bayesian Image Restoration with two Applications in Spatial Statistics (with discussion). *Annals of the Institute of Statistical Mathematics,* 43, 1-59.

Brezger, A., Lang, S. (2006). Generalized Structured Additive Regression based on Bayesian P splines. *Computational Statistics and Data Analysis* 50, 967-991.

Craddock, S., 2004, Beyond Epidemiology: Locating AIDS in Africa, in HIV and AIDS in Africa, Kalipeni et al., Blackwell, Oxfordemic, Journal of the American Medical Association, 270(17): 2083-2086.

Fagbamigbe, A.F., Adebayo, S.B. & Idemudia, E. (2016). Marital status and HIV prevalence among women in Nigeria: Ingredients for evidence-based programming. *Int.l Journal. of Infectious Diseases*, 48: 57 – 63.

Federal Ministry of Health. (2008). National HIV/AIDS and Reproductive Health and Serological Survey, 2007 (NARHS Plus). Abuja Nigeria: Federal Ministry of Health.

Federal Ministry of Health. (2013). National HIV/AIDS and Reproductive Health and Serological Survey, 2012 (NARHS Plus). Abuja, Nigeria: Federal Ministry of Health.

Gayawan, E., Adebayo, S. B., Ipinyomi,R. A. & Oyejola, B. A. (2010). Modelling fertility curves in Africa. *Demographic Research* 22(10): 211-236.

Hastie, T., Tibshirani, R., 1990. Generalized additive models. Chapman and Hall, London

Johnson K, Way A (2006) Risk factors for HIV infection in a national adult population: evidence from the 2003 Kenya Demographic and Health Survey. Journal of Acquired Immune Deficiency Syndromes 42(5): 627–636.

Joint United Nations Programme on HIV/AIDS and World Health Organization Working Group on Global HIV/AIDS and STI Surveillance. Guidelines for measuring national HIV prevalence in population-based surveys. Geneva: WHO/UNAIDS; 2005.

Lang, S., Brezger, A. (2004). Bayesian P-Splines *Journal. of Computational and Graphical Statistics*, 13, 183-212.

MacQueen, K., 1994, The Epidemiology of HIV Transmission: Trends, Structure and Dynamics, Annual Review of Anthropology, 23 : 509-526.

Manda SO, Lombard CJ, Mosala T (2012) Divergent spatial patterns in the prevalence of the human immunodeficiency virus (HIV) and syphilis in South African pregnant women. Geospatial health 6(2): 221–231.

Mishra, V., Montana, L., Neuman, M. (2007). Spatial modelling of HIV prevalence in Kenya. In Demographic and Health Research.

Ngesa, O., Mwambi, H. & Achia, T. (2014). Bayesian Spatial Semi-Parametric Modelling of HIV Variation in Kenya. PLoS ONE 9(7): e103299. https://doi.org/10.1371/journal.pone.0103299.

Rue, H., Held, L. (2005). Gaussian Markov random fields. Theory and applications. Chapman & Hall/CRC, Boca Raton- London-NewYork-Singapore.

Saphonn V, Hor LB, Ly SP, Chhuon S, Saidel T, Detels R. (2002). How well do antenatal clinic (ANC) attendees represent the general population? A comparison of HIV prevalence from ANC sentinel surveillance sites with a population-based survey of women aged 15–49 in Cambodia. *Int. Journal of Epidemiology*; 31:449-55.

World Health Organization and Joint United Nations Programme on HIV/AIDS. Reconciling antenatal clinic-based surveillance and population-based survey estimates of HIV prevalence in sub Saharan Africa. Geneva: WHO and UNAIDS; 2003.

Zaba, B., Carpenter, L.M., Boerma, J.T., Gregson, S., Nakiyingi, J. & Urassa, M. (2000). Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa. *AIDS*;14:2741-50.

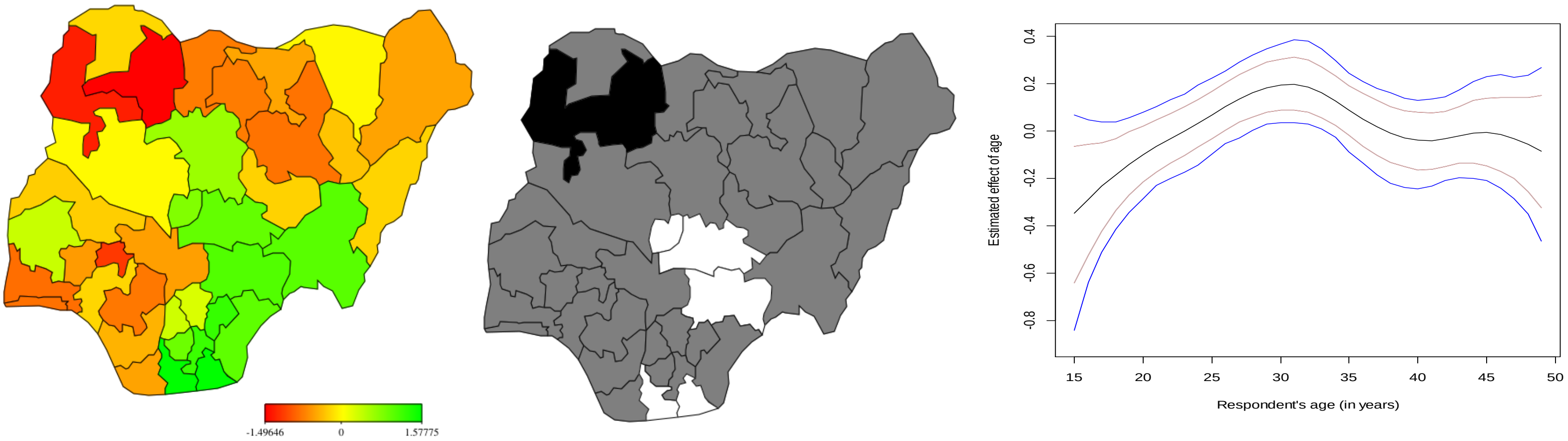

Figure 1: (a) and (b) Maps of Nigeria showing the total (structured and unstructured) spatial effects and the map of significance. For map of significance, dark showing negative credible interval and white showing positive credible interval. (c) is the nonlinear effect of respondent's age. Included are the 95% (in blue) and 80% (grey) confidence intervals.

Table 1: Posterior Odds Ratio and 95% credible interval for three models

Variables	**a) Model with un-split spatial effect**			**b) Model without spatial effects**			**c) Model with structured & unstructured spatial effects**		
	Posterior mean	**95% Cr. Interval**		**Posterior mean**	**95% Cr. Interval**		**Posterior mean**	**95% Cr. Interval**	
		Lower	**Upper**		**Lower**	**Upper**		**Lower**	**Upper**
Constant	0.035	0.000	481.473	0.035	0.023	0.054	0.000	0.000	0.006
Year of study	0.836	0.679	1.048	0.782	0.640	0.951	0.794	0.643	0.969
Qur'anic	0.951	0.627	1.490	0.865	0.506	1.386	0.836	0.489	1.359
Primary	0.959	0.702	1.355	2.104	1.596	2.742	2.009	1.518	2.643
Secondary	0.913	0.658	1.269	2.125	1.608	2.847	1.883	1.397	2.522
Higher	0.667	0.458	1.006	1.607	1.076	2.374	1.342	0.900	2.048
Urban	0.925	0.710	1.187	0.942	0.752	1.194	1.102	0.870	1.383
Poorer	1.132	0.826	1.558	0.898	0.666	1.190	0.928	0.688	1.244
Average	1.557	1.124	2.115	1.232	0.920	1.644	1.312	0.984	1.747
Richer	1.584	1.126	2.235	1.239	0.877	1.707	1.326	0.954	1.832
Richest	1.711	1.142	2.562	1.301	0.910	1.831	1.342	0.946	1.915
Currently Married	1.228	0.863	1.734	0.887	0.668	1.200	1.134	0.787	1.610
Formerly Married	0.990	0.472	1.925	1.992	1.371	2.943	2.094	1.391	3.121
Transactional Sex	1.151	0.819	1.567	1.091	0.791	1.493	1.067	0.773	1.456
Had multiple sexual partners	1.202	0.975	1.495	1.297	0.996	1.694	1.171	0.883	1.532
Current sexual activities	1.137	0.811	1.633		0.811	1.633	0.791	0.600	1.045
Non-marital sexual practices	0.772	0.561	1.060	0.772	0.561	1.060	1.638	1.153	2.268
Sexual debut at < 15 Years	0.568	0.000	266.459	0.568	0.000	266.459	106.994	3.197	17,537.8
Can't remember age at sexual debut	0.599	0.000	287.324	0.599	0.000	287.324	104.221	3.267	17,245.9
Sexual debut at >= 15 Years	0.471	0.000	224.589	0.471	0.000	224.589	118.079	3.746	18,787.9
Experienced STIs in the last 3 months	1.002	1.000	1.003	1.002	1.000	1.003	0.942	0.942	0.942
Comprehensive knowledge of HIV	1.000	0.999	1.000	1.000	0.999	1.000	1.000	1.000	1.000