Spatial analysis of factors associated with HIV infection in sub-Saharan Africa using DHS

Muluye Desta & Monica Magadi
Social Research Methodology Centre,
Department of Sociology,
City University London

Much has been written about the depth and extended effect of HIV/AIDS on Sub-Saharan Africa's societal fabric. However, with the availability of recent data and refined methodology of analysis, the continuity of HIV/AIDS research is essential for informed and organized engagement at the affected community level. According to the 2008 UNAIDS report, the disproportionate nature of the HIV/AIDS pandemic is alarming in that sub-Saharan Africa – about two-thirds (67%) of all people living with HIV in the world are in sub-Saharan Africa (SSA). The situation is much bleak when these figures are viewed at country level as most of the severely affected countries are from the sub-Saharan African region.

As medical science has not yet achieved a level to produce a vaccine or effective medication for preventing or treating HIV/AIDS, the alternative research – looking for ways to curb risk taking behaviour through proxy factors, has been a priority for researchers in other disciplines. The 2008 UNAIDS report emphasizes this: "...HIV is a biological entity that is responsive to medical interventions, but the epidemic has continued to expand, largely due to the failure to tackle societal conditions that increase HIV risk and vulnerability"(UNAIDS, 2008). It is, therefore, indispensable to examine the factors that influence the society in order to closely understand the biological, psycho-social and socio-cultural ways that can help curb the spread of HIV/AIIDS pandemic at the affected community level.

This paper is part of a wider study on HIV/AIDS in sub-Saharan Africa and has three fold objectives: (i) to examine spatial variation of HIV infection through a relatively new geo-additive modelling technique; (ii) to identify factors associated with spatial variation between females and males and across national borders.

We use a relatively new approach - spatial modelling with a Bayesian posterior predictive distribution that takes into account all sources of uncertainty in the model formulation while at the same time incorporating spatial variation. The Bayesian inference is based on the

'Markov Neighbourhood' and 'Posterior probability distribution' carried out using one of the latest simulation techniques - Markov Chain Monte Carlo (MCMC).

The additional advantage of MCMC simulation technique is that it uses a 'Markov property' that conditional probability distribution of future states of a process, given the present state, depends only upon the current state, that is, it is conditionally independent of past states given the present state. A typical Bayesian geo-additive model based on MCMC regression technique further uses Markov random field also called a 'prior' and two-dimensional P-splines (where a prior is a direct generalization of a first order random walk to two-dimensions and the P-splines approach assumes that an unknown smooth function of a covariate can be approximated by a polynomial spline of degree one defined by a set of equally spaced knots).

In this paper, we use a Bayesian generalized additive mixed model with a logit link function based on Markov random field priors (Fahrmeir and Lang, 2001; Smith and Kohn, 1996; Kammann and Wand, 2003). Geo-additive mixed models allow for the analysis of the effects of categorical and continuous covariates as well as regional (spatial) effects within a unified semi-parametric Bayesian framework for inference. The logit model for the probability that a respondent in a region S is HIV positive can be modelled as:

$$\eta_i = f_1(X_{i1}) + f_2(X_{i2}) + ... + f_k(X_{ik}) + f_1(W_{1j}) + f_2(W_{2j}) + f_{sp}(S_i)$$

where $f_1(X_{i1}) + f_2(X_{i2}) + ... + f_k(X_{ik})$ are categorical predictors for individual *i*;

$f_1(W_{1j}) + f_2(W_{2j})$ are nonlinear/continuous factors and

$f_{sp}(S_i)$ is the effect of the spatial covariate S_i.

An important advantage of spatial analysis is that it usually appeals to special considerations not found in a standard statistical analysis. These considerations include the fact that a map has a common visual appeal to anyone without advanced knowledge to interpret or understand the results of a statistical modelling. Hence, despite the advanced nature of the underlying principles of spatial modelling, the outcome has the extra advantage due to its suitability for policy briefs as it can be conveyed to anyone at any level. Besides the visual perception of the spatial distribution of a phenomenon (say disease), spatial analysis is also very useful to translate prevailing patterns into both objective and quantifiable outcome.

Recent development in data collection has enabled the collection of actual HIV test data along with a cross sectional survey on demographic and health records. This new technique of data collection has been pioneered by Macro International Inc. in collaboration with developing countries where such surveys are conducted. The innovative approach of HIV test data is collected alongside demographic and health issue interviews where an interviewer collects a 'dried blood spot sample on a filter paper card using blood from a finger prick with a single-use, spring-loaded, sterile lancet labelled with a barcode that links with the consenting respondent's house hold and individual line number'. The HIV test result is then merged with other socio-demographic data by first removing information that could potentially identify the individual respondent.

We use data from the Demographic and Health Surveys (DHS) conducted by 19 sub-Saharan African countries conducted since 2001. The main outcome variable is HIV infection and the covariates included in the models include background socio-demographic factors, sexual behaviour factors, awareness factors and household headship as well as poverty indicator factors.

Model building was carried out for males, females and the combined sub-Sahara Africa data by including the above factors step-by-step to see the spatial and effect variation. The findings include, among others, that the inclusion of 'sexual behaviour factors' such as risky sexual behaviour (with no condom and with non-spousal partner), multiple partners, pre-marital sex, age at first sexual activity, circumcision entails a sudden change in the odds of HIV infection (an increase in the odds ratio compared to other factors). On the other hand, inclusion of 'awareness factors' such as media exposure, HIV awareness, education and knowledge of someone with HIV brings a change in the models as a trade-off (balancing effect) to the sexual behaviour effect.

The geo-additive model residuals were then mapped to see the spatial variation when the above group of covariates were included into the model. The result provides significant cross national spatial variation where the introduction of some of the covariates brings distinct changes on the HIV infection incidence map particularly in Southern and Eastern African countries.

Results indicate that the spatial variation of HIV infection becomes noticeable when sexual behaviour factors: ‘risky sexual engagement’, ‘premarital sex’, ‘multiple sexual partners’, ‘age at sexual activity’, and ‘marital status’ are controlled for in the spatial model. The results also revealed emerging regions in Western Africa moving toward mid-to-high range of HIV infection.