Applications of geospatial analysis to surveillance data: a spatial examination of HIV/AIDS prevalence in Zambia

Imelda K. Moise · Ezekiel Kalipeni

Abstract Techniques of spatial statistics and GIS are applied to socio-economic, demographic and HIV sentinel data to characterize the geographical distribution of HIV prevalence in Zambia and to estimate current prevalence rates. Maps of the 4 years under study (i.e. 1994, 1998, 2002 and 2004) reveal a spatial variation in HIV prevalence with urban and provincial districts having higher prevalence than rural areas. However, there was an overall trend of decrease in HIV prevalence across the country. The year 2004 exhibited the most reduction, coinciding with protective sexual behavior campaigns operating in the country. Risk factors for HIV prevalence included literacy rates, unemployment, poverty and urban residency. Analysis of regression residual maps indicated high spatial autocorrelation: Moran's I value was 0.28 ($z = 4.12$ and $p < 0.000$) and a highly significant Robust LM (lag; $p < 0.00000$) suggestive of contagious and hierarchic spatial diffusion processes. High HIV prevalence rates among interdependent districts and locational similarity patterns suggest that HIV control programs in the country would require an integrated approach combining HIV prevention messages as well as an understanding of social and cultural interactions between interdependent districts that produce behavioral diffusion of HIV prevalence rates.

Keywords HIV and AIDS · Geospatial analysis · Surveillance data · Zambia

I. K. Moise (✉)
Department of Geography, University of Illinois at Urbana-Champaign, Urbana, IL 61801-3671, USA
e-mail: imoise2@illinois.edu

E. Kalipeni
National Science Foundation, 4201 Wilson Boulevard, Arlington, VA 22230, USA
e-mail: ekalipen@nsf.gov

Introduction

Human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) rank among the most important public health problems in Zambia with an estimated 1.2 million people living with HIV and 630,000 with AIDS (WHO 2000; Zambia Central Statistical Office et al. 2002). Since its discovery in Zambia over two decades ago (Chanda and Gosnell 2006; Fylkesnes et al. 1994; Garbus 2003; WHO 2005; WHO 2005), the virus has continued to be a major cause of chronic infections and premature deaths (UNAIDS/WHO 2006). It has decreased life expectancy and is a serious impediment to economic development. The current estimate of HIV prevalence among adults in Zambia is about 15.2%, twice the average prevalence in Sub-Saharan Africa. Like many other infectious diseases, HIV/AIDS prevalence varies among subpopulations. For example, the

prevalence of HIV seropositivity is 40% in tuberculosis patients, 30–40% in STD patients, 10–30% in prisoners, 10% in antenatal women and only 5–10% for blood donors (Tembo and Hira 1989). Although current estimates indicate that HIV/AIDS prevalence is stabilizing or even declining at the national level (Syacumpi et al. 2003; UNAIDS/WHO 2005; WHO 2005), there is concern that these estimates, may be under- or over-predictions of the magnitude of the disease. Moreover these estimates are based on national rather than local data. Consequently, there is continued interest in improving techniques for estimating local rates of HIV/AIDS.

To this end, this paper has two main objectives. The first objective is to examine the spatial pattern of HIV prevalence in Zambia by utilizing local scanty sentinel data obtained from testing pregnant women attending designated antenatal clinics. In so doing, this study uses geographic information systems (GIS) to identify the patterns of HIV rates in the country over time and generate mean HIV prevalence rates per district. The second objective is to enter the generated rates for the year 2004 into a spatial statistics software package, namely, GeoDa (Anselin 2003, 2005; GeoDa Center for Geospatial Analysis and Computation 2010a, b) to generate two kinds of regression models (ordinary least squares and spatial lag regressions) that explain the observed spatial variations in HIV/AIDS rates using several socioeconomic independent variables measured at district level. The resultant regression residuals are then subjected to further tests to check for spatial autocorrelation, an integral part of the mapping technique which we used, i.e. the inverse distance weighted interpolation (IDW). However, when the variations of the generated mean HIV/AIDS rates are explained by a set of independent variables in a multiple regression model, the presence of spatial autocorrelation in the residuals is a strong indication of the fact that there are other independent variables that were not included but should have been included in the selected set of independent variables. Although the disciplines of spatial analysis and GIS have developed quite independently of one another (Robinson 2004), together not only do these techniques provide insight into the HIV epidemic but can also aid in mapping risk areas, identify causal factors driving HIV transmission in the country as well as aid in decision making and surveillance.

Literature review

Over the past 25 years a substantial number of studies have been conducted focusing on various aspects of the HIV/AIDS epidemic. Many of these studies have concentrated on the biomedical aspects of the disease with the aim of finding a cure. However, a select few have concentrated on methods of how to better estimate the magnitude of the disease in terms of its spread over space and among certain demographic groups. Other studies have attempted to explain the factors which facilitate its spread over time and space, particularly in the hardest hit areas of southern and eastern Africa. Below we offer a brief literature review in terms of studies that attempt to better estimate the rates and offer explanatory factors for its rapid spread.

Statistical approaches to estimation of HIV/AIDS rates

Various statistical approaches have been used in estimating HIV/AIDS prevalence. The three commonly used ones are: (1) backcalculation techniques, which attempt to reconstruct the number of AIDS cases over time to past trends in HIV infection (e.g. Brookmeyer and Damiano 1989; Brookmeyer and Gail 1988, 1986, 1994; Kaplan and Brookmeyer 1999), (2) extrapolation (Karon et al. 1989; Morgan and Curran 1986), which attempts to fit a model to the HIV incidence curve and extrapolate into the future, and (3) HIV dynamic transmission models, which attempt to compare contact patterns within regions or risk groups (Anderson et al. 1986; Bailey 1975; Garnett et al. 2004; Gregson et al. 2002; Morris and Kretzschmar 1997). However, the successful use of these models for HIV/AIDS estimates in Africa has been limited, partly because of inadequate means by which HIV/AIDS data can be collected and monitored. Monitoring difficulties coupled with limited resources often present challenges in obtaining reliable information on the prevalence of HIV/AIDS by region (Kalipeni and Zulu 2008; Salomon and Murray 2001) and, as a result, there has been debate over the validity and reliability of current HIV prevalence estimates. Much of what is known about HIV prevalence in Zambia is based on data that is obtained from pregnant women attending designated antenatal clinics.

Recent efforts to model and estimate HIV/AIDS prevalence have utilized predominantly HIV sentinel data, population-based HIV survey data and mathematical models or have relied on comparisons that were somewhat limited spatially. For example, in Sub-Saharan Africa, a maximum likelihood approach was applied to a mathematical model for the estimation of HIV incidence, prevalence and AIDS mortality over time (Salomon and Murray 2001). In Zambia extrapolation was used to demonstrate a correlation between HIV sentinel data and data for the general population (Fylkesnes et al. 1998). A similar study in Zambia revealed a significant decline in HIV prevalence (Fylkesnes et al. 2001). Research is needed that can combine sero-surveillance data with socio-economic, demographic and cultural variables, analyzing them for significant relationships, and investigating what geographical factors may lead to increased risk of infection, and how and why certain locations have higher HIV prevalence rates than others.

GIS combined with methods of spatial statistics provides powerful new tools for understanding the epidemiology of infectious diseases (Chadee and Kitron 1999; McLafferty and Cromley 2002; Robinson 2004). GIS offers several advantages over traditional methods of spatial analysis and provides a robust approach to the study of HIV/AIDS "as it has the ability to store, manipulate, and display data linked to locations" (Bentley-Condit and Hare 2006; Kitron 1998). Spatial analysis, on the other hand, can be used to extract the spatial relationship embedded between geographic locations explicitly allowing for the description and testing of spatial patterns among the geographic locations (Anselin 1988).

Factors affecting the spread of HIV

A number of studies have developed conceptual models rooted in the contributions of social science to medicine and public health in order to understand the multiple dimensions of vulnerability to HIV/AIDS in Africa (Kalipeni et al. 2007; Yeboah 2007). The frameworks attempt to identify the major drivers of vulnerability to HIV/AIDS in Africa. Oppong (1998), in applying vulnerability theory to Ghana's HIV/AIDS situation, argues that while all human beings are biologically susceptible to infection by different diseases such as HIV/AIDS, certain social and economic factors place some individuals and social groups in situations of increased vulnerability. For example, social and economic conditions that require people to leave their spouses for long periods (for example the military and migrant workers), increases the risk of casual sexual liaisons and the related vulnerability to HIV. Thus vulnerability is multidimensional and results from multiple combinations of factors.

It is indeed gratifying to see that an increasing number of scholars are asking the questions of what and how factors converge to create risky conditions for particular individuals and groups within sub-Saharan Africa. Rather than focusing on patterns of sexual exchange per se, critical social science research investigates the multiple and interrelated reasons why women or men get placed in situations that increase their likelihood of engaging in risky behaviors, thereby increasing their vulnerability to HIV. One of the primary factors increasing HIV transmission in southern Africa, for example, is a pattern of male out-migration driven by depressed national economies on the one hand, and on the other hand labor-intensive mining industries in South Africa (Adepoju 2003; Girdler-Brown 1998; Decosas et al. 1995; Brockerhoff and Biddlecom 1999; Haour-Kinpe and Rector 1996; Upton 2003; Coffee et al. 2007). The result is male workers away from their families for long periods of time and women increasingly unable to fend for themselves while their husbands are gone (Chirwa1998; Campbell 1997; Zuma et al. 2003). Other studies on the impact of migration on the proliferation of HIV/AIDS have also found similar results as those in Africa (Hong et al. 2006). In other areas, prolonged war has disrupted local economies, displaced populations, and rendered longstanding social practices untenable (Bond and Vincent 1997; Kalipeni and Oppong 1998; Lyons 2004). In most areas of sub-Saharan Africa, economies are struggling and poverty is rampant, making it difficult for families to stay together in one place or for individuals to maintain viable incomes without resorting to potentially risky sexual economies (Akeroyd 1997; Addai 1999; Becker et al. 1999; Decosas1996; Fourn and Ducic 1996; Rugalema 2004; Lurie et al. 2004).

In this vein, Eileen Stillwaggon's work on the determinants of the spread of HIV maintains that the over sexualized African is a myth, instead the

variation in space of other social, economic, gender and health determinants can explain the distribution of HIV prevalence across sub-Saharan Africa (see Stillwaggon 2000, 2001, 2002, 2003). In light of the above brief literature review, there are several areas of inquiry that need to be addressed in a comprehensive manner in order to understand the complex macro and micro-level drivers of HIV (see Ackerman and de Klerk 2002; Kalipeni et al. 2004; Kalipeni et al. 2007). In their exhaustive literature reviews of the socioeconomic factors that fuel the rapid spread of HIV/AIDS in southern Africa, Nyindo (2005) and Kalipeni et al. (2007) identify six major drivers which constitute the major multiple and overarching dimensions of vulnerability. These are the historical context of colonialism and labor migration, gender, poverty and disease burden, global forces, culture, and government attitudes. As noted earlier, the central aim of this paper is to conduct a geographically grounded study of the temporal-spatial spread of HIV/AIDS in Zambia. Given data constraints, the study utilizes some of the socioeconomic determinants of HIV/AIDS rates highlighted in the brief review of the literature above. Specifically, we conduct multiple regression analysis of the 2004 dependent variable, HIV/AIDS rates, against a set of independent variables which included literacy rates,
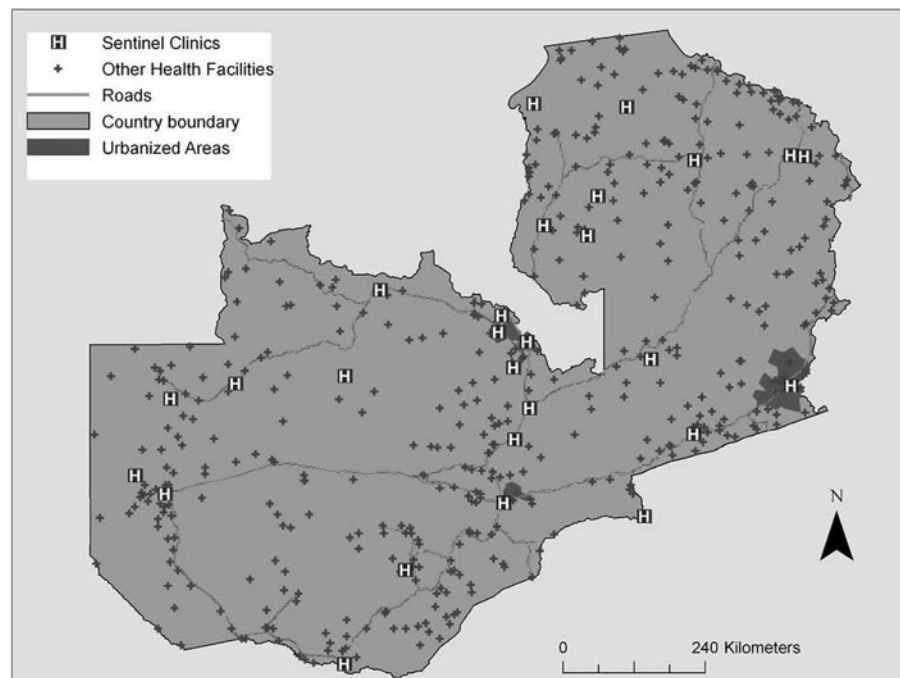
unemployment rates, population urban, and population poor for the year 2004 at district level.

Data and methods

HIV prevalence data

Sentinel surveillance data used for the maps came from United Nations Development Program (UNDP 2005), Zambia Vulnerability Assessment Control (VAC 2005), National AIDS Committee (NAC) annual reports and Centra Technology Inc. Data obtained from Centra Technology covered the 26 sentinel clinics that are designated for screening pregnant women attending antennal clinics in Zambia (Fig. 1). Originally Centra Technology obtained the dataset from the United States Bureau of the Census, HIV surveillance database, and geocoded it by adding centroids (x, y coordinates; Kalipeni and Zulu 2008; Turnock 2001). These data provide the annual proportion of pregnant women attendees that tested positive for HIV infection in the country. However, there was considerable variation in the number of data points that were reported for each consecutive year at each clinic as clinics started testing HIV/AIDS at different times. As a result, there were

Fig. 1 Spatial distribution of Zambia's 568 health facilities, 27 sentinel clinics and major roads. *Source*: Authors, data obtained from SAHIMS, http://www.sahims.net/

insufficient data points for some clinics, which resulted in the selection of 4 years with sufficient data for further analysis, namely 1994, 1998, 2002 and 2004 for the study. For the analysis in this paper we used district as the geographic unit of data aggregation ($N = 72$). As such we obtained additional district level independent variables from other

sources such as UNDP and VAC annual reports (Fig. 2). Spatial data for health facilities (point data) and political boundary (district) shapefiles were downloaded from the Southern Africa Humanitarian and Disaster Management (SAHIMS) GIS library at: (http://www.sahims.net/gis/GIS_library). Demographic and socio-economic data such as population,

Fig. 2 Administrative units of Zambia's 72 districts and urbanized areas. *Source*: Authors, data obtained from SAHIMS, http://www.sahims.net/

poverty status, unemployment rates and life expectancy were obtained from the Zambia Central Statistics Censuses and other important data sources such UNDP. Due to limitations of data availability, the socioeconomic data were obtained only for the year 2004.

Data interpolation analysis (inverse distance weighting)

A database was created that allowed for computation and visualization of patterns from the data. Included in the database were sentinel data, number of pregnant women sampled, their age, and HIV rates per year, socio-economic and demographic data, road network and political boundary shapefiles. A deterministic technique, the inverse distance weighted interpolation (IDW), under the spatial analyst extension in Arc-GIS (Booth and Crosier 2004) was used to spatially interpolate HIV surveillance data into continuous smooth maps based on the 4-year district-level HIV prevalence data. IDW was chosen over kriging due to the small sample size of the dataset. IDW has been used in this context for a long time (Beck et al. 2002; Bohra and Andrianasolo 2001; Carrat and Valleron 1992; Kalipeni and Zulu 2008; Webster et al. 1994). A recent study by Kalipeni and Zulu (2008) confirms the fact that IDW offers better results when the number of points is small in comparison to kriging, hence our preference in using IDW for this study.

The following parameters were used: power value of 2 and inclusion of 8 surrounding data points. A spatial resolution of 1 km (km) × 1 km or 0.08333 decimal degrees was used for the output raster image. A spatial mask for Zambia was used to limit the interpolation within the Zambian political boundary. The result was a Zambian map for each year that was composed of 1 km × 1 km pixels, each containing an estimate of the HIV rate for that specific pixel. We then used these interpolated values to generate mean HIV estimates per year by district, taking into account the data from all available data points.

Spatial autocorrelation analysis

Spatial statistics are helpful for describing the spatial variations exhibited by the response process of a given phenomenon (Anselin and Getis 1992). In this study, local Moran's I statistics (LISA) is calculated

since the interest is in identifying specific observations or districts that exhibit spatial autocorrelation with their neighbors. Two tests exist in GeoDa to do this: Moran's I, and Lagrange Multiplier test. In this study, Moran's I statistics (Anselin 1988) are used to identify specific observations or districts that exhibit spatial autocorrelation with their neighbors. Moran's I has been extended to the diagnosis of spatial dependence in the presence of covariates (Cliff and Ord 1973). Moran's I measures autocorrelation in regression residuals.

Moran's I was calculated as:

$$I = \frac{N \sum_i \sum_j w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\left(\sum_i \sum_j w_{ij}\right) \sum_i (X_i - \bar{X})^2} \tag{1}$$

where N is the number of areas (districts), i refers to a particular district with j referring to that district's neighbors, w_{ij} is an element of a row-standardized weight spatial matrix (i.e., elements of a row sum to 1) corresponding to the observation pair i and j. The sample mean is subtracted from X_i and X_j to calculate deviations from the sample mean for areas i and j.

The expected value of Moran's I under the null hypothesis is $-1/(N - 1)$. If observed Moran's I is greater than expected, neighboring districts have similar incidence rates of HIV prevalence. If observed Moran's I is smaller than expected, neighboring districts have dissimilar rates.

Since spatial dependence was identified via local tests of spatial autocorrelation given that spatial heterogeneity in parameters did not account fully for the observed univariate spatial dependence, the next step involved modeling this spatial autocorrelation via covariates (independent variables). It is important that we define what we mean by spatial dependence at the outset. Spatial dependence exists when the value associated with one location is dependent on those of other locations (see GeoDa Center for Geospatial Analysis and Computation 2010a, b). This can result from spatial interaction effects (e.g., externalities or spill-over effects) or from measurement error (e.g. related to a mismatch between the scale at which a phenomenon occurs and how it is measured; GeoDa Center for Geospatial Analysis and Computation 2010a, b). Since it is quite challenging to distinguish whether a location is impacted by its neighboring values or is just different, it can be difficult to separate spatial dependence from spatial

heterogeneity. On the other hand a spatial lag is a variable that essentially averages the neighboring values of a location (the value of each neighboring location is multiplied by the spatial weight and then the products are summed). Spatial lags are used in the computation of global and local Moran's I, as well as in spatial lag (Wy) and spatial error models (We). They can also be computed as separate variables (e.g., WX) in GeoDa (GeoDa Center for Geospatial Analysis and Computation 2010a, b).
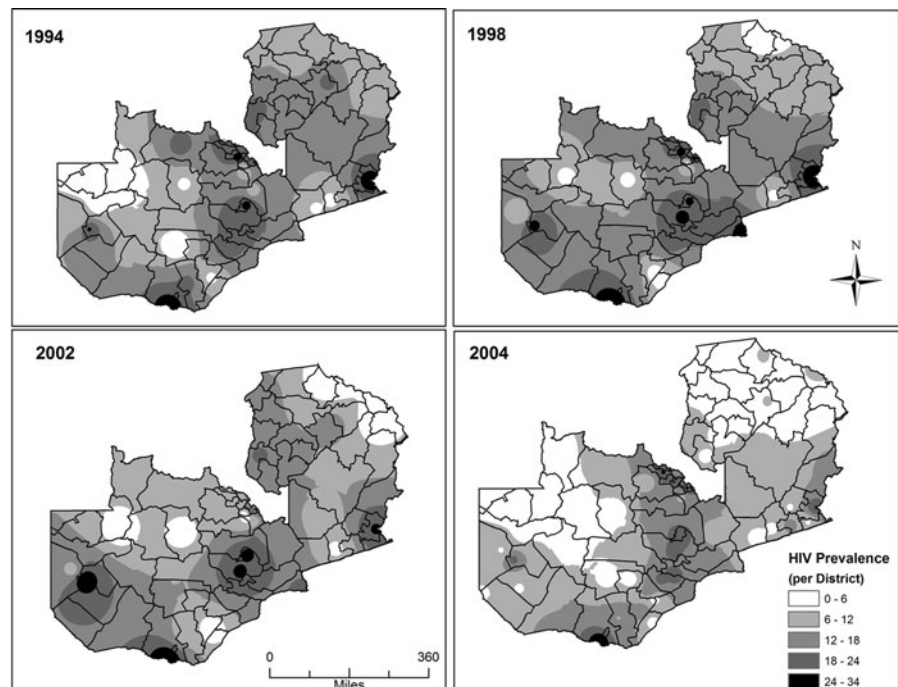
A classical ordinary least square (OLS) regression was run with diagnostics for spatial lag to determine whether the covariates observed fully modeled the observed spatial dependence. Diagnostics for spatial dependence indicated spatial lag dependence in the presence of covariates, evidence consistent with a spatial contagious diffusion process.

Results

Spatial patterns of HIV prevalence rates (inverse distance weighted analysis)

Smoothed maps generated from Inverse Distance Weighting for the years 1994, 1998, 2002 and 2004 reveal a spatial variation in HIV prevalence rates

(Fig. 3). Urban areas had higher prevalence rates than rural areas. Districts with a high prevalence rate of over 20% included Ndola and Kitwe on the Copperbelt; Lusaka, Luangwa and Kafue in Lusaka province; Livingstone and Mazabuka in Southern Province; Solwezi in Northwestern province (1994 and 1998); Kabwe and Kapiri Mposhi in Central province; Chipata in eastern province and Mongu in Western province. Most of the districts with high prevalence rates also straddle major roads and truck routes (Fig. 4). These results also conform to a 2001/2002 demographic health survey that reported urban residents to be twice as likely to be infected with HIV as rural residents.

HIV prevalence was highest in 1994 (Fig. 3) and was more distinct in urban areas and in provincial headquarter districts than in rural areas. Though prevalence was slightly lower in 1998 with areas of high incidence in northern and northwestern districts disappearing, the size of the areas with elevated rates seems to have increased, especially in Lusaka, eastern and southern districts. Prevalence rates continued to decline into 2002 and 2004. In 2002, prevalence was higher in western, central and in two eastern districts with a new elevated area appearing in Luapula's provincial district of Mansa. By 2004, the decrease is more prominent for both rural and urban

Fig. 3 IDW generated maps of spatial patterns of HIV prevalence rates for 1994, 1998, 2002, 2004

Fig. 4 HIV prevalence in 2004 in relation to major roads and truck routes
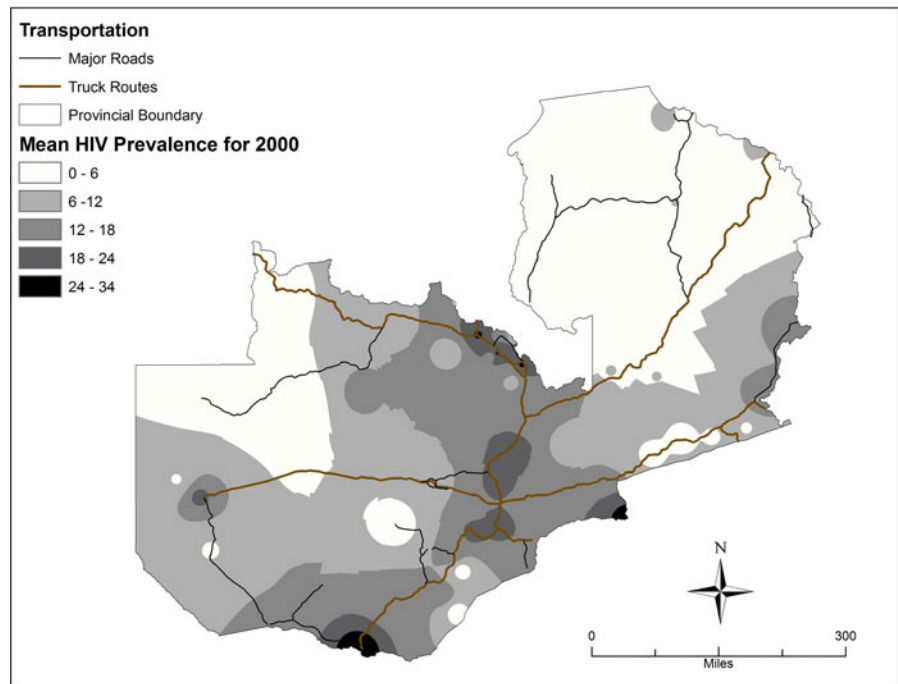
areas; however, prevalence among Copperbelt districts seems to be increasing when compared to 2002. The nature of the observed spatial patterns (declining rates) is reflective of protective sexual behavior campaigns that have been operating in the country over the past two decades. This could also indicate that there are fewer new infections than the number of people who are dying from AIDS related deaths.

Analysis of mean average prevalence rates calculated at the district level still show evidence of higher rates in urban areas. It can be seen for the 1994 spatial pattern that higher rates are spread across the country with more west–east and north-central trends (Fig. 5). On the other hand, the year 1998 saw continued higher rates in southern and central areas with less intensity in the northern region. The 2002 map shows a decrease in rates among southern, eastern and Copperbelt districts with an intense central region. The year 2004 saw a decrease in northern and North-western districts, however, districts in the south, central and Copperbelt provinces still have high HIV prevalence rates. As expected, HIV prevalence in urban areas remains high in all the 4 years, making urban areas epicenters for the disease with HIV prevalence ranging between 24 and 34%.

However, as noted above, between 1994 and 2004, there were dramatic declines in HIV/AIDS prevalence rates for many districts as shown in Fig. 6. Out of the 72 districts only 15 experienced increases in their HIV rates ranging from 3.08% increase for Siavonga district to 150% increase for Luangwa districts. All the 15 districts that experienced increases over the 10 year period (1994–2004) were largely rural districts. On the other hand, the majority of the districts (57 in total) experienced declines ranging from 1 to 81% declines. Districts with major urban centers and provincial headquarters experienced the largest declines in HIV rates. This is also testament to the fact that urban areas were the first to experience the highest rates while in later years the disease began to diffuse to surrounding rural areas in a contagious process.

Socio-economic variables associated with HIV prevalence

Results obtained from the initial OLS model detected a significant relationship between literacy rates, urban residence, unemployment, poverty and HIV prevalence, implying that an increase would induce an increase in HIV prevalence in the district. An

Fig. 5 Maps of mean HIV
prevalence rates by district
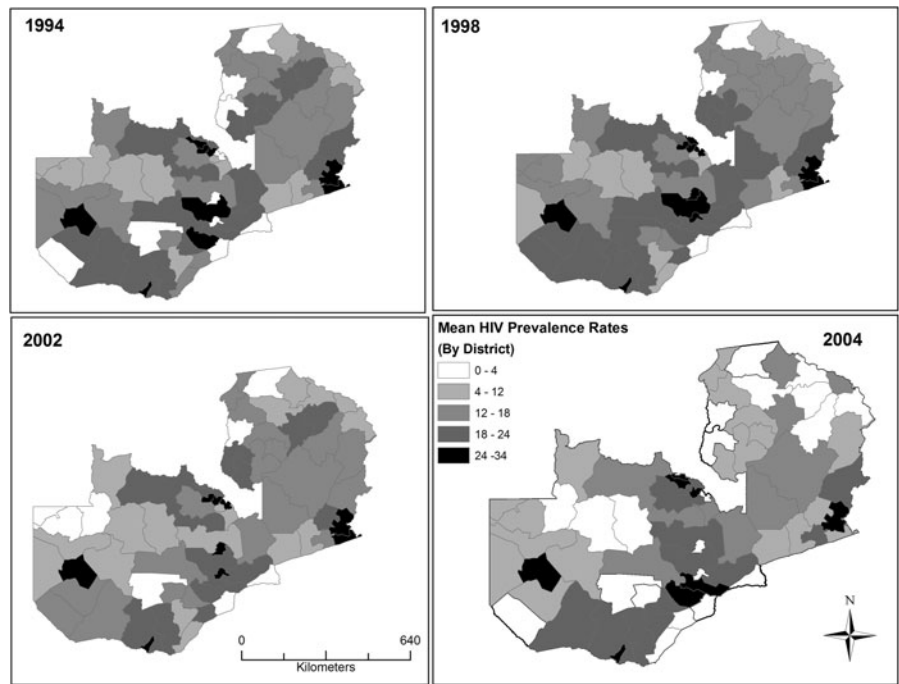for the years 1994, 1998,
2002, 2004

Fig. 6 Percent change
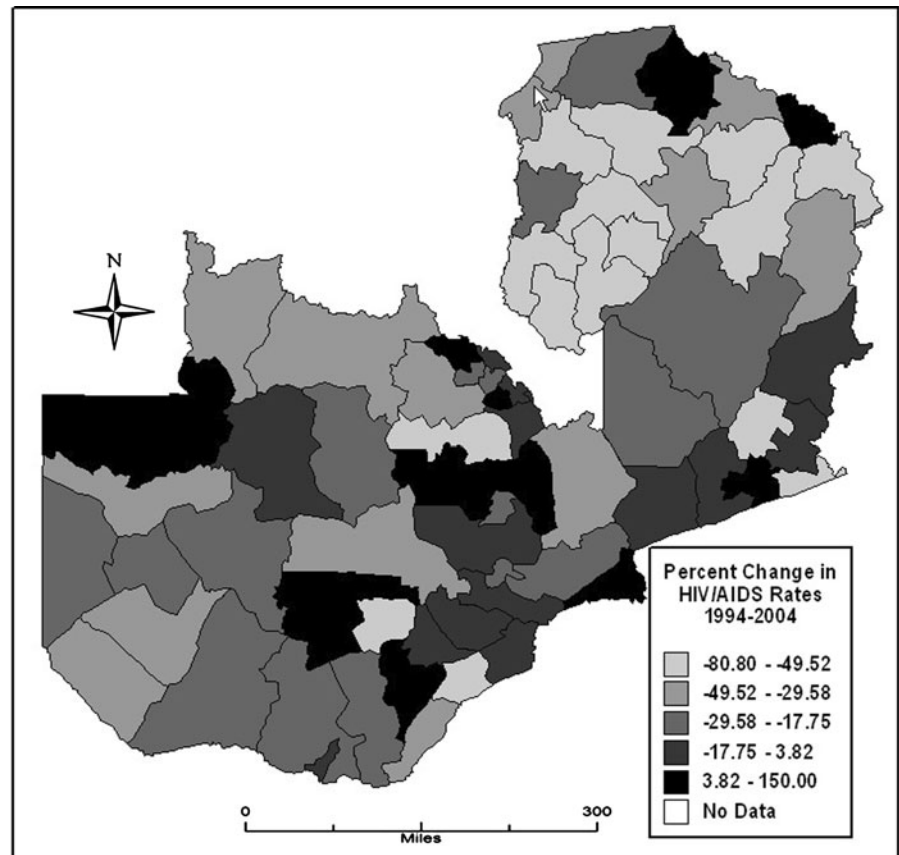in HIV/AIDS rates,
1994–2004

Table 1 Results of OLS regression model

Variable	Coefficient	Std error	t-statistic	Probability
Constant	−0.201	1.415	−0.142	0.000
Literacy rates >25	0.112	3.268	0.021	0.021
Unemployment rates	0.182	0.089	2.043	0.016
Population urban	4.243	1.511	2.808	0.006
Population poor, 2003	6.836	2.337	2.925	0.011

Log likelihood: −225

AIC: 463.9

R-square: 0.62

SC: 476.9

increase in the number of people living in an urban area will result in an increase in that areas' HIV prevalence rate. This result is also in agreement with the hypothesis that urbanization increases people's risks for HIV infection. Rural residence has a negative value suggesting that districts that have less population would have lower rates of HIV prevalence (Table 1).

A significantly positive Moran's I value was observed, 0.28 ($z = 4.12$ and $p < 0.000$) for the 2004 HIV prevalence data indicating that districts sharing a border are more similar with respect to HIV prevalence than would be expected under the null hypothesis of complete spatial randomness (Table 2). Furthermore, both Lagrange Multiplier (lag) and Lagrange Multiplier (error) tests on the OLS model are highly significant ($p < 0.000$ and $p < 0.000$). And of the robust form, Robust LM (lag) is more significant than Robust LM (error; $p < 0.00000$ compared to $p < 0.009$). Both tests suggest that there is spatial dependence and the spatial lag model

Table 2 Diagnostics for spatial dependence from OLS model

Test	MI/DF	Value	Prob
Moran's I (error) 0.286419	4.12	0.0000378	0.0000000
Lagrange Multiplier (lag)	1.00	32.5063660	0.0000000
Robust LM (lag)	1.00	26.8532791	0.0000002
Lagrange Multiplier (error)	1.00	12.3788027	0.0004342
Robust LM (error)	1.00	6.7257158	0.0095033
Lagrange Multiplier (SARMA)	2.00	39.2320818	0.0000000

should be used in the analysis because the Robust LM (lag) has a lower p-value. In short the results of OLS regression analysis suggest a statistically significant relationship between HIV prevalence for 2004 and literacy rates for those aged 15 and 25, unemployment, urban residence and poverty. Diagnostics for spatial dependence from the OLS model suggest that spatial dependence exists and that we can proceed with confidence to use the spatial lag model for spatial analysis.

Spatial dependence occurs when a value observed in one location depends on the values observed at neighboring locations (Anselin 2003). The spatial lag reflects observations that are not independent. The observed spatial lag is suggestive of a possible diffusion process as HIV prevalence in one district predicts an increased likelihood of similar rates in neighboring places. This also implies that HIV prevalence in Zambia is highly social in nature, and understanding the interactions between interdependent districts is critical to understanding the rates of HIV infection.

Table 3 depicts the results of the spatial lag model for the association of HIV prevalence for 2004 and explanatory variables. Though there are some minor differences in the significance of the other regression coefficients between the spatial lag model and the OLS model, rural residence is more significant than before ($p < 0.04$), but more importantly, the significance of literacy rates, urban residence and unemployment rates changes from significant to insignificant ($p < 0.26$, $p > 0.33$, $p < 0.17$), respectively.

Table 3 Results of spatial LAG model

| Variable	Coefficient	Std error	z-Value	Probability
W_PREV_2004	0.783	0.067	11.537	0.000
Constant	4.553	1.516	3.002	0.002
Literacy rates >25	−0.028	0.025	−1.131	0.258
Unemployment rates	0.089	0.065	1.379	0.167
Population urban	1.051	1.099	0.955	0.339
Population poor, 2003	3.382	1.693	1.997	0.045

Log likelihood: −203.8

AIC: 421.6

R-square: 0.72

SC: 437.5

The spatial autoregressive coefficient (W_PREV_2004) is estimated as 0.78, and is highly significant ($p < 0.000000$). Checking the order of the Wald (W), Likelihood ratio (LR), and Lagrange Multiplier (LM) statistics on the spatial autoregressive coefficient, it was found that $W = 132.25$ (the square of the z-value of the asymptotic t-test), $LR = 43.7$, and $LM = 32.5$ (from OLS model). This corresponds to the expected order ($W > LR > LM$; Anselin 2005) and suggests that this Maximum Likelihood Estimation (MLE) is good. Also, it can be seen from the changes of the p-values of the variables that the explanatory power of these variables that was attributed to their in-district

value, was due to the neighboring locations. This is picked up by the coefficient of the spatially lagged dependent variable (W_PREV_2004).

When the two model residual maps are compared both the spatial lag model and the OLS model (Fig. 7) show evidence of spatial autocorrelation. However, the spatial lag map shows that the spatial patterns of HIV prevalence in Zambia are somewhat different to the OLS model. In both models, however, the location of large residuals is the same (darkest shading). Both maps do suggest that similarly colored areas tend to be in similar locations, which could be positive autocorrelation (Moran's I test for residual

Fig. 7 Maps of regression residuals for the OLS and LAG models
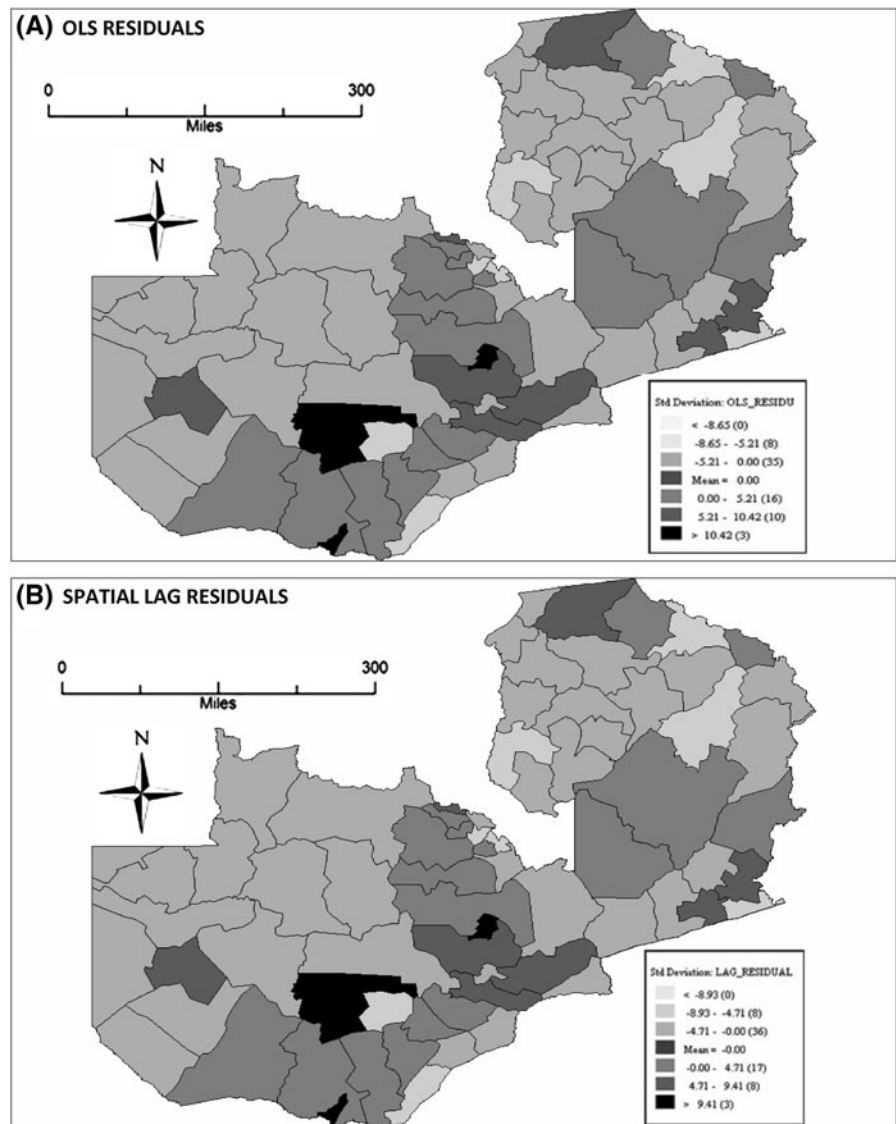
spatial autocorrelation is positive and highly significant). There is a tendency to over-predict (negative residuals) in the outlying areas and a tendency to under-predict (positive residuals) in the core, suggesting the possible presence of heterogeneity (Anselin 2005). A negative standard deviation means that the predicted values exceeded the actual values. Both maps exhibit areas of spatially autocorrelated residuals. This was an unexpected result which indicates a need to further control for spatial autocorrelation for the residuals. It shows that there are other significant factors at play which led to the spatial clustering pattern of the residuals. We may therefore need to look for other independent variables in later studies for which the spatial error model will be needed.

Discussion

Spatial patterns of HIV prevalence rates (IDW)

Analyses of spatial patterns of HIV for the 4 years under study (1994, 1998, 2002 and 2004) reveal a gradual declining trend in HIV prevalence across the country in geographic space. This result holds whether spatial pattern is assessed using inverse distance weights or by mean HIV prevalence obtained from zonal statistics as illustrated by Figs. 3 and 4. It can be seen from both IDW and mean HIV prevalence maps that HIV prevalence is highest in 1994 and is more pronounced in urban areas and in provincial head-quarter districts than in rural areas for all 4 years.

One possible explanation of the observed high prevalence in urban and provincial districts could be due to high population density in these areas as well as the existence of most services. The result has been a rural/urban drift coupled with high unemployment in urban areas and perhaps indicating higher levels of risky sexual behaviors. The high rate of rural/urban drift may have contributed to high HIV prevalence rates in these areas since urban areas have the highest prevalence rates. Unfortunately, the patterns and number of those moving to cities is unknown. Nevertheless, we strongly suggest that responsible sexual behavior messages should be targeted to both rural and urban residents so that as people migrate from rural to urban areas, they would already have had the information needed to protect themselves from contracting the deadly disease.

Provincial districts exhibiting high prevalence rates are points of interest and they include Mongu in Western, Solwezi in North-Western, Livingstone in Southern, and Luangwa and Chipata in Eastern province. These areas provide an opportunity to examine the role played by social amenities that exist in urban areas versus rural areas. Clearly, most socio-economic exchanges in Zambia take place in provincial districts. Provincial districts tend to have political influence, good infrastructure as well as social amenities such as nightclubs, hospitals and banks. As a result, people from other districts come to these areas every month-end to collect their salaries giving them an opportunity to engage in more leisure activities including drinking, bar hopping and sexual relations. Thus, high prevalence rates in these areas may have resulted from human interaction within and among the provincial districts.

Additionally, with an exception of Mongu, most of the provincial districts border other neighboring countries indicating that factors other than socio-economic indicators need to be considered in order to explain the source of this clustering in prevalence rates. One mechanism that could explain areas of high incidence at borders is human migration from neighboring countries and people visiting these areas from elsewhere. For example, Livingstone is known to be a tourist destination. As a result, people visiting the city could be responsible for the high prevalence rates in the area as well as the rise in commercial sex work. Similarly, Luangwa district is home to one of Zambia's largest wild life reserves which attract thousands of visitors per year.

Spatial association of HIV prevalence

Further analysis of regression residual maps suggests strong spatial autocorrelation of HIV prevalence in the country, as one would expect on account of its strong link with socio-economic and demographic variables. There is a tendency to over-predict (negative residuals) in the outlying areas and a tendency to under-predict (positive residuals) in the core suggesting the possible presence of heterogeneity (Anselin 2005) which holds with both OLS and spatial lag models.

The geographical distribution of HIV prevalence in the country deviates significantly from a complete random process as contiguous districts form zones of

high prevalence rates, suggesting that HIV prevalence in one district predicts an increased likelihood of similar rates in neighboring places. These results suggest that one district's proximity to its neighbors might be linked to the diffusion of HIV between neighboring districts. This is based on the premise that location and distance are important forces at work. In addition, districts of high HIV prevalence rates may be related to social activities that affect one's chances of indulging in risky behaviors. Figure 7a, b show residual maps that depict this observed spatial autocorrelation.

The spatially autocorrelated areas of high prevalence obtained from IDW and mean HIV prevalence rates in conjunction with OLS and spatial lag model residual maps are reason to investigate other underlying explanations for the geographical distribution of HIV prevalence in Zambia. Most importantly, the relationship between the 2004 HIV prevalence and socio-economic variables may elucidate the role played by socio-economic and demographic determinants in HIV rate variations.

A positive significant association between the 2004 HIV prevalence and literacy rates is evidenced by the results of the OLS model, which contradicts the hypothesis that literacy rate for those aged 15 and 25 years is negatively associated with HIV prevalence in the country. The assumption was that, the more educated people are the more likely they will be in seeking HIV/AIDS information thus a reduction in HIV prevalence. The result suggests on the contrary that the more educated one is, the more likely they are to engage in risky behaviors putting themselves at high risk of contracting the virus. This may also be ascribed to the fact that educated people tend to have high earning power as well as more disposable income to support risky behaviors.

Additionally, the association between literacy rate and HIV prevalence can also be explained by the socio-economic differences among different genders as a result of educational attainment. For example, more men in Zambia are educated when compared to women and are more likely to be mobile, working from an urban environment to a rural environment and vice versa. As such men tend to have more social networks and opportunities for high paying jobs as well as leisure activities that lead to extra marital affairs. Extra marital affairs have become an ethos for those living in cities in Zambia.

From these results, it appears that education and its financial benefits have become a risk factor in HIV transmission (UNAIDS 1998). This is consistent with previous studies that showed that HIV transmission is directly linked to education, urban living and high income (Gisselquist et al. 2003; Potts 2003). This study shows that a person's socio-economic status can reinforce a person's behavior that can put s/he at risk for contracting HIV. If this is true, HIV control programs in the country would require an integrated approach combining HIV prevention messages as well as an understanding of social and cultural interactions between interdependent districts that produce behavioral diffusion of HIV prevalence rates.

The significant negative association between rural residence and HIV prevalence supports the hypothesis that HIV prevalence is lower in rural areas than in urban areas. This is likely to be for various cultural and socio-economic reasons. First, Zambia has been experiencing a rural/urban drift since its independence in 1964. Mostly, young energetic people, groups that are more likely to engage in risky sexual behaviors, migrate to urban areas in search of jobs leaving rural areas with a lower-risk adult population. Secondly, rates could be lower in rural areas mainly because of cultural and local norms. While illicit sexual behavior may be tolerated in cities, in most rural areas it is seen as a sign of lack of good morals that brings shame on ones' family. As such people tend to refrain from indulging in sexual relations for fear of embarrassment and being stigmatized. Additionally, a smaller population in terms of numbers also means that people tend to know their neighbors very well, every person's affairs are everybody else's business as such there is no room for secrecy or unfaithfulness.

Poverty is also correlated significantly with HIV prevalence. As poverty level increases in a district so are that district's HIV prevalence rates. This was expected as studies have shown a link between poverty and HIV transmission in different countries in Sub-Saharan Africa (UNAIDS 1998; Whiteside 2002). Similarly, women and young girls have been reported to engage in commercial sex work for economic reasons (Whiteside 2002; Zambia Central Statistical Office 2002). This in itself raises concern considering the high number of youths in the country who have finished high school and even college but

are lacking jobs. If nothing is done to alleviate this situation, the observed declining rates might be reversed. HIV prevention messages alone might not be enough to fight the deadly disease that is robbing the country of its young talented professionals. There is need for immediate government intervention that directly relates to intervening in ending unemployment among youths.

Conclusion

This paper uses GIS to model the spatial distribution of HIV prevalence rates in Zambia from a scanty point surveillance data set for the years 1994, 1998, 2002, and 2004. We illustrate the power of geography to generate HIV/AIDS rates at any geographic unit, in this case at the district level. We then further use advanced spatial statistical techniques in GeoDa to analyze the presence or absence of spatial autocorrelation in space using a set of independent variables that help explain the spatial variation of HIV/AIDS in Zambia. The results confirm the fact that HIV/AIDS rate are very high in urban areas and provincial headquarter districts. We conclude by noting that these results can be validated with more refined data such as that currently being collected by demographic health surveys. Data on HIV prevalence among different age groups as well as locales would be extremely useful and provide a more nuanced GIS analysis to validate the findings in this paper. In addition to investigating prevalence in different sub populations, it would be useful to evaluate the role of human migration within and in between districts. Analysis of the patterns of human migration in relation to HIV transmission would assist in explaining the high prevalence rates in urban and provincial districts. Specifically, the nature and scale of high HIV/AIDS prevalence rates around border districts would better be explained using both international and country level migration patterns.

References

Ackerman, L., & de Klerk, G. W. (2002). Social factors that make South African women vulnerable to HIV infection. *Health Care for Women International, 23*, 163–172.

Addai, I. (1999). Ethnicity and sexual behavior. *Social Biology, 46*(1–2), 17–32.

Adepoju, A. (2003). Continuity and changing configurations of migration to and from the Republic of South Africa. *International Migration, 41*(1), 3–28.

Akeroyd, A. (1997). Sociocultural aspects of AIDS in Africa: Occupational and gender issues. In G. C. Bond, J. Kreniske, I. Susser, & J. Vincent (Eds.), *AIDS in Africa and the Caribbean* (pp. 11–30). Boulder, CO: Westview Press.

Anderson, R. M., Medley, G. F., May, R. M., & Johnson, A. M. (1986). A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *Mathematical Medicine and Biology, 3*, 229–263.

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Boston: Kluwer.

Anselin, L. (2003). *GeoDa 0.9 user's guide*. Urbana-Champaign: Spatial Analysis Laboratory. (Urbana-Champaign: Department of Geography, University of Illinois).

Anselin, L. (2005). *Exploring spatial data with GeoDaTM: A workbook*. Urbana-Champaign: Spatial Analysis Laboratory, Department of Geography, University of Illinois at Urbana-Champaign.

Anselin, L., & Getis, A. (1992). Spatial statistical analysis and geographic information systems. *The Annals of Regional Science, 26*, 19–33.

Bailey, N. T. J. (1975). *The mathematical theory of infectious diseases and its applications*. London: Griffin.

Beck, L. R., Kitron, U., & Bobo, M. (2002). Remote sensing, GIS, and spatial statistics: Powerful tools for landscape epidemiology. In W. J. M. Martens & A. J. McMichael (Eds.), *Health impacts of global environmental change: Concepts and methods*. New York: Cambridge University Press.

Becker, C., Dozon, J.-P., Obbo, C., & Toure, M. (Eds.). (1999). *Experiencing and understanding AIDS in Africa*. Dakar: Codesria, Karthala, IRD.

Bentley-Condit, V. K., & Hare, T. S. (2006). Using geographic information systems and spatial statistics to examine the spatial dimension of animal social behaviour: A baboon example. *Journal of Mathematics Sciences, 1*(1). http://www.scientificjournals.org/articles/1028.htm.

Bohra, A., & Andrianasolo, H. (2001). Application of GIS in modeling of dengue risk based on sociocultural data: Case of Jalore, Rajasthan, India. *Dengue Bulletin, 25*, 92–102.

Bond, G., & Vincent, J. (1997). AIDS in Uganda: The first decade. In G. C. Bond, J. Kreniske, I. Susser, & J. Vincent (Eds.), *AIDS in Africa and the Caribbean*. Boulder, CO: Westview Press.

Booth, B., & Crosier, S. (2004). *ArcGIS 9: Getting started with ArcGIS. GIS by ESRI*. Redlands, California: ESRI.

Brockerhoff, M., & Biddlecom, A. E. (1999). Migration, sexual behavior and the risk of HIV in Kenya. *International Migration Review, 33*(4), 833–857.

Brookmeyer, R., & Damiano, A. (1989). Statistical methods for short-term predictions of the AIDS incidence. *Statistics in Medicine, 8*, 23–34.

Brookmeyer, R., & Gail, M. H. (1986). Minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United States. *Lancet, 2*(8519), 1320–1322.

Brookmeyer, R., & Gail, M. H. (1988). A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. *Journal American Statistical Association, 83*, 301–308.

Brookmeyer, R., & Gail, M. H. (1994). *Monographs in epidemiology and biostatistics*. New York: Oxford University Press.

Campbell, C. (1997). Migrancy, masculine identities and AIDS: The psychosocial context of HIV transmission on the South African Gold Mines. *Social Sciences and Medicine, 45*, 273–283.

Carrat, F., & Valleron, A. J. (1992). Epidemiologic mapping using the "kriging" method: Application to an influenza-like epidemic in France. *American Journal of Epidemiology, 135*, 1293–1300.

Chadee, D. D., & Kitron, U. (1999). Spatial and temporal patterns of imported malaria cases and local transmission in Trinidad. *American Journal of Tropical Medicine and Hygiene, 61*, 513–517.

Chanda, D., & Gosnell, D. (2006). The impact of tuberculosis on Zambia and the nursing workforce. *Journal of Issues in Nursing, 11*(1). http://www.openj-gate.org/Articlelist.asp?Source=1&Journal_ID=80280.

Chirwa, W. (1998). Aliens and AIDS in southern Africa: The Malawi-South Africa debate. *African Affairs, 97*, 53–79.

Cliff, A. D., & Ord, J. K. (1973). *Spatial autocorrelation*. London: Pion Press.

Coffee, M., Lurie, M. N., & Garnett, G. P. (2007). Epidemiology and social modelling the impact of migration on the HIV epidemic in South Africa. *AIDS, 21*(3), 343–350.

Decosas, J. (1996). HIV and development. *AIDS, 10*(Suppl 3), S69–S74.

Decosas, J., Kane, F., Anarfi, J. K., Sodji, K. D. R., & Wagner, H. U. (1995). Migration and AIDS. *Lancet, 36*(8978), 826–828.

Fourn, L., & Ducic, S. (1996). Epidemiological portrait of acquired immunodeficiency syndrome and its implications in Benin. *Sante, 6*(6), 371–376.

Fylkesnes, K., Brunberg, H., & Msiska, R. (1994). *The socio-economic impact of AIDS: The current HIV/AIDS situation and future demographic impact*. Lusaka, Zambia: National AIDS/STD/TB and Leprosy Programme Document, Ministry of Health, Zambia.

Fylkesnes, K., Musonda, R. M., Sichone, M., Ndhlovu, Z., Tembo, F., & Monze, M. (2001). Declining HIV prevalence and risk behaviors in Zambia: Evidence from surveillance and population-based surveys. *AIDS, 15*, 907–916.

Fylkesnes, K., Ndhlovu, Z., Kasumba, K., Mubanga Musonda, R., & Sichone, M. (1998). Studying dynamics of the HIV epidemic: Population-based data compared with sentinel surveillance in Zambia. *AIDS, 12*, 1227–1234.

Garbus, L. (2003). *Country AIDS policy analysis project: HIV/AIDS in Zambia*. San Francisco: AIDS Policy Research Center, University of California.

Garnett, G. P., Grassly, N. C., Boerma, J. T., & Ghys, P. D. (2004). Maximizing the global use of HIV surveillance data through the development and sharing of analytical tools. *Sexually Transmitted Infections, 80*(Supplement 1), i1–i4.

GeoDa Center for Geospatial Analysis and Computation. (2010a). *GeoDa and methods for geospatial analysis*. Arizona State University. http://geodacenter.asu.edu/.

GeoDa Center for Geospatial Analysis and Computation. (2010b). *Glossary of Key Terms*. Arizona State University. Available online at: http://geodacenter.asu.edu/node/390.

Girdler-Brown, B. (1998). Eastern and southern Africa. *International Migration, 36*(4), 513–551.

Gisselquist, D., Potterat, J., Brody, S., & Vachon, F. (2003). Let it be sexual: How health care transmission of AIDS in Africa was ignored. *International Journal of STD and AIDS, 14*(3), 148–161.

Gregson, S., Nyamukapa, C., Garnett, G., Mason, P., Zhuwau, T., Caraël, M., et al. (2002). Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet, 359*, 1896–1903. doi:10.1016/S0140-6736(02)08780-9.

Haour-Kinpe, M., & Rector, R. (1996). *Crossing borders: Migration, ethnicity and AIDS*. Bristol, PA: Taylor and Francis.

Hong, Y., Stanton, B., Li, X., Yang, H., Lin, D., Fang, X., et al. (2006). Rural-to- urban migrants and the HIV epidemic in China. *AIDS and Behavior, 10*, 421–430.

Kalipeni, E., Craddock, S., Oppong, J. R., & Ghosh, J. (Eds.). (2004). *HIV/AIDS in Africa: Beyond epidemiology*. Oxford: Blackwell Publishers.

Kalipeni, E., Ghosh, J., & Mkandawire-Valhmu, L. (2007). The multiple dimensions of vulnerability to HIV/AIDS in Africa: A social science perspective. In M. Mundi (Ed.), *Women, AIDS and access to health care in sub-Saharan Africa: Approaches from the social sciences* (pp. 45–60). Barcelona: Medicus Mundi Catalunya.

Kalipeni, E., & Oppong, J. (1998). The refugee crisis in Africa and implications for health and disease: A political ecology approach. *Social Science and Medicine, 46*(12), 1637–1653.

Kalipeni, E., & Zulu, L. (2008). Using GIS to model and forecast HIV/AIDS rates in Africa, 1986–2010. *Profession Geographer, 60*(1), 33–53.

Kaplan, E. H., & Brookmeyer, R. (1999). Snapshot estimates of recent HIV incidence rates. *Operations Research, 47*(1), 29–37.

Karon, J. M., Devine, O. J., & Morgan, W. M. (1989). *Predicting AIDS incidence by extrapolating from recent trends*. New York: Springer.

Kitron, U. (1998). Landscape ecology and epidemiology of vector-borne diseases: Tools for spatial analysis. *Journal of Medical Entomology, 35*(4), 435–445.

Lurie, P., Hintzen, P. C., & Lowe, R. A. (2004). Socioeconomic obstacles to HIV prevention and treatment in developing countries: The roles of the International Monetary Fund and the World Bank. In E. Kalipeni, S. Craddock, J. Oppong, & J. Ghosh (Eds.), *HIV/AIDS in Africa: Beyond epidemiology* (pp. 204–212). Malden, MA: Blackwell Publishers.

Lyons, M. (2004). Mobile populations and HIV/AIDS in East Africa. In E. Kalipeni, S. Craddock, J. Oppong, & J. Ghosh (Eds.), *HIV/AIDS in Africa: Beyond epidemiology* (pp. 175–190). Malden, MA: Blackwell Publishers.

McLafferty, L. S., & Cromley, E. K. (2002). *GIS and public health*. New York: The Guilford Press.

Morgan, W. J., & Curran, J. W. (1986). Acquired immunodeficiency syndrome: Current and future trends. *Public Health Reports, 101*, 459–465.

Morris, M., & Kretzschmar, M. (1997). Concurrent partnerships and the spread of HIV. *AIDS, 11*, 641–648.

Nyindo, M. (2005). Complementary factors contributing to the rapid spread of HIV-I in sub-Saharan Africa: A review. *East African Medical Journal, 82*(1), 40–46.

Oppong, J. (1998). A vulnerability interpretation of the geography of HIV/AIDS in Ghana, 1986–1995. *The Professional Geographer, 50*(4), 437–448.

Potts, A. (2003). *Socio-economic status and HIV/AIDS in sub-Saharan Africa. Behavioral science and health*. University of Utah. http://www.bsh.utah.edu/modules.php?op=modload&name=News&file=article&sid=6&mode=thread&order=0&thold=0. Accessed on 5 May 2007.

Robinson, T. P. (2004). Spatial statistics and geographical information systems in epidemiology and public health. *Parasitology, 47*, 83–127.

Rugalema, G. (2004). Understanding the African HIV epidemic: An appraisal of lay discourse on HIV/AIDS in Africa. In E. Kalipeni, S. Craddock, J. Oppong, & J. Ghosh (Eds.), *HIV/AIDS in Africa: Beyond epidemiology* (pp. 191–203). Malden, MA: Blackwell.

Salomon, J. A., & Murray, C. J. L. (2001). Modeling HIV/AIDS epidemics in sub-Saharan Africa using seroprevalence data from antenatal clinics. *Bulletin of World Health Organization, 79*, 596–607.

Stillwaggon, E. (2000). HIV transmission in Latin America: Comparison with Africa and policy implications. *South African Journal of Economics, 68*(5), 985–1011.

Stillwaggon, E. (2001). AIDS and poverty in Africa. *Nation, 272*(20), 22–25. Online resource at http://www.thenation.com/doc.mhtml?i=20010521&c=2&s=stillwaggon.

Stillwaggon, E. (2002). HIV/AIDS in Africa: Fertile terrain. *Journal of Development Studies, 38*(6), 1–22.

Stillwaggon, E. (2003). Racial metaphors: Interpreting sex and AIDS in Africa. *Development and Change, 34*(5), 809–832.

Syacumpi, M. M., Liywali, K., Mbale, M., & Syacumpi, M. (2003). *Country analysis of family planning and HIV/AIDS: Zambia*. Washington DC: Futures Group International, Policy Project (USAID Contract No. HRN-C-00-00-00006-00). http://www.policyproject.com/pubs/countryreports/Zam_FPHIV.pdf.

Tembo, G., & Hira, S. (1989). An epidemiological review of HIV in Zambia. *International Conference on AIDS, 5*, 1001. June 4–9; (Abstract No. M.G.P.23), Montreal, Quebec. http://gateway.nlm.nih.gov/MeetingAbstracts/102181221.html. Accessed 5 May 2007.

Turnock, B. J. (2001). *Public health: What it is and how it works*. Sudbury, MA: Jones and Bartlett.

UNAIDS. (1998). *Report on the global HIV/AIDS epidemic*. Geneva: UNAIDS.

UNAIDS/WHO. (2005). *Update: Epidemiological fact sheets on HIV/AIDS and sexually transmitted infections*. Zambia. Geneva: UNAIDS/WHO.

UNAIDS/WHO. (2006). *Report on the global AIDS epidemic*. Geneva: UNAIDS/WHO.

United Nations Development Programme. (2005). Zambia. United Nations Development Programme website. http://www.undp.org.zm/. Accessed 5 June 2005.

Upton, R. L. (2003). 'Women have no tribe': Connecting carework, gender, and migration in an era of HIV/AIDS in Botswana. *Gender & Society, 17*(2), 314–322.

Webster, R., Oliver, M. A., Muir, K. R., & Mann, J. R. (1994). Kriging the local risk of a rare disease from a register of diagnoses. *Geographical Analysis, 26*, 168–185.

Whiteside, A. (2002). Poverty and HIV/AIDS in Africa. *Third World Quarterly, 23*(2), 313–332.

WHO. (2000). *Guidelines for second generation HIV surveillance*. Geneva: World Health Organization and UNAIDS.

WHO. (2005). *Zambia: Summary country profile for HIV treatment scale-up*. Geneva: World Health Organization. http://www.who.int/hiv/HIVCP_ZMB.pdf. Accessed on 5 May 2007.

Yeboah, I. E. A. (2007). HIV/AIDS and the construction of Sub-Saharan Africa: Heuristic lessons from the social sciences for policy. *Social Science and Medicine, 64*(5), 1128–1151.

Zambia Central Statistical Office, Zambia Central Board of Health, ORC Macro MEASURE DHS+. (2002). *Zambia demographic and health survey, 2001–2002. Preliminary report*. Lusaka, Zambia: Central Statistical Office.

Zambia Vulnerability Assessment Committee (ZVAC). (2005). *2005 Vulnerability and needs assessment*. Lusaka: The Zambia Vulnerability Assessment Committee. http://www.sarpn.org.za/documents/d0001309/P1547-June2005_VAC_Zambia.pdf. Accessed 5 June 2007.

Zuma, K., Gouws, E., Williams, B., & Lurie, M. (2003). Risk factors for HIV infection among women in Carletonville, South Africa: Migration, demography and sexually transmitted diseases. *International Journal of STD and AIDS, 14*(12), 814–817.