Geographic Distribution of Human Immunodeficiency Virus in South Africa

Immo Kleinschmidt[*],
Tropical Epidemiology Group, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom, Telephone: 44 20 7927 2103, Fax: 44 20 7636 8739, Immo.kleinschmidt@lshtm.ac.uk.

Audrey Pettifor,
Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27514, a.pettifor@rhrujhb.co.za.

Natashia Morris,
Medical Research Council of South Africa, 491 Ridge Road, Durban, South Africa, Telephone: 27-31-203-4806, Fax: 27-31-203-4704, nmorris@mrc.ac.za.

Catherine MacPhail, and
Reproductive Health and HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa, c.macphail@rhrujhb.co.za

Helen Rees
Reproductive Health and HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa, h.rees@rhrujhb.co.za.

Abstract

In common with most countries, little is know about the geographic distribution of human immunodeficiency virus (HIV) in South Africa. Variations in HIV infection of persons 15–24 years of age were modeled and mapped using generalized linear spatial models and Bayesian prediction based on data from a national HIV household survey conducted in 2003 and comprising 11,904 youth from approximately 700 enumeration areas that were randomly selected from the national census. The maps show considerable variation in HIV prevalence within provinces. The lowest levels were found in inland rural areas of the Western Cape, and the highest in northwestern parts of KwaZulu Natal, southern Mpumalanga, and eastern Free State. Prevalence of HIV was associated with ethnicity, urban status, and unemployment. Detailed maps of HIV prevalence can be effectively used in guiding and focusing intervention programs to areas of particular need.

INTRODUCTION

The prevalence of infection with human immunodeficiency virus (HIV) in South Africa has been continuing on a significant upward trend.[1,2] Young people, and particularly young women, have been identified as being especially at risk of infection.[3–7] Nationally representative HIV seroprevalence surveys are an important means of assessing whether HIV prevention campaigns are having an impact on the prevalence of infection in successive

[*]Address correspondence to London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. Immo.kleinschmidt@lshtm.ac.uk.

cohorts of young people reaching the age of sexual maturity. Understanding the variation of HIV infection within a country is essential for determining where prevention and treatment programs need to focus resources.

Prevalence of HIV infection in South Africa, as elsewhere, has previously been reported either as a national or a provincial average.[1,2] Such aggregate estimates may eclipse the geographic heterogeneity in the HIV epidemic in a country. The HIV seroprevalence in a local area is often only conjecture, based on hearsay or on non-representative samples of the population. For more effective monitoring of local HIV prevalence, it is necessary to report the results of national HIV prevalence surveys in a geographically disaggregate form while avoiding the reporting or mapping of prevalence levels that may be subject to large random errors caused by the small samples on which they are based.

We have used data from the 2003 South African national household youth survey of HIV infection and sexual behavior[5] to derive spatial statistical models of HIV prevalence in 15–24 year-old men and women separately and, using these models, performed spatial prediction of HIV prevalence for the entire area of South Africa. We also derived the prediction uncertainty at each map location. We therefore present the first continuous HIV prevalence maps for South Africa. The simultaneous mapping of estimated prevalence and the confidence interval of the prediction map ensures that HIV prevalence maps do not assume a false claim of precision for any particular locality. Nevertheless, detailed mapping of HIV prevalence in South Africa provides us with a picture of the geographic variation of the epidemic that overall national (or provincial) HIV prevalence estimates are unable to do. Such maps can be used to further guide local prevention and care efforts to ensure their relevance to the local HIV situation.

METHODS

This study was based on data collected in a national household survey of HIV infection and sexual behaviors of young people 15–24 years of age conducted in 2003.[5] Using the 2001 national census[8] as a sampling frame, 714 enumeration areas (EAs) were randomly selected, stratified by province and enumeration area type. Within each EA, one eligible youth was randomly selected from each sampled household. Selected individuals were asked to answer questions relating to HIV risk behavior and to provide a specimen of oral fluid to test for antibodies to HIV.[9] All HIV testing was anonymous, but linked to questionnaires. The methodology and results of this survey have previously been reported in greater detail.[5,10]

For the purpose of spatial modeling, the total number of male and female respondents was aggregated for each of the 697 sampled EAs that could be geo-located, and the proportion infected with HIV, by sex, was calculated. Using geographic information systems,[11] the latitude and longitude of the centroid of each sampled EA was determined (Figure 1). The map surface of South Africa was divided into 1,945 pixels of approximately 25 km × 25 km, for which model predictions of HIV prevalence were made that excluded areas for which no census data apply, such as nature reserves and game parks. Socioeconomic and demographic data were extracted from the 2001 census for each of the sampled EAs and for each of the pixels. The value for each pixel was calculated by summing the variable values over the census areas making up the area of the pixel, and allocating values in proportion to the area of the census entity contained within the pixel.

Area related possible explanatory variables were first tested in non-spatial univariate logistic regression models to determine their potential for being significantly associated with HIV infection in 15–24 year-old persons. Variables that were not associated with HIV infection in non-spatial models, by sex, were not further explored. The following census variables

were investigated: proportion of residents who were black Africans, proportion of an area urban,[12] proportion of 20–64 year-old persons unemployed, population density, proportion of households that were informal, proportion of 15–19 year-old persons who did not finish school, proportion of households without electricity, proportion of 20–24 year-old persons not living with partners, proportion of heads of households less than 19 years of age, and proportion of 16–24 year-old persons with a deceased mother. From the limited number of variables available from the census, all those were selected that had been shown to be linked to HIV infection in the survey, from other literature, or speculatively as potential explanatory variables.

Generalized geostatistical spatial models using Markov Chain Monte Carlo (MCMC) simulation were fitted to HIV prevalence at sampled locations (EAs) to identify significant covariates from the census. Bayesian kriging was used to perform spatial prediction of HIV prevalence at the grid of all map locations (pixels). Details of the method are given in the Appendix. Spatial modeling and spatial prediction was carried out using the package geoRglm in the statistical software system R.[13]

RESULTS

Spatial analysis and mapping were carried out separately for women and for men because of the large sex difference in HIV prevalence. The HIV test results of a total sample of 6,148 females and 5,610 males from the 697 EAs were used in this analysis. Analysis and mapping of HIV prevalence for women was based on 688 EAs at which the EA sample included women. Analysis and mapping of prevalence in men was based on a different but overlapping set of 688 EAs in which men were included in the EA sample (there were nine EAs without women, and an additional nine without men). The mean sample size for each EA was 8.8 for women and 8.0 for men. The mean HIV prevalence in EA samples was 13.6% for women (range = 0–100%, median =10%). The observed mean prevalence was 4.8% in men at the EA level (range = 0–100%, median = 0%). Figure 1 shows the location of the 697 EAs on which the modeling and prediction was based.

In spatial logistic regression models, the following area variables were significantly associated with HIV infection in 15–24 year-old women: proportion of residents who were black Africans, proportion of 20–64 year-old women who were unemployed in an area, proportion of households that were informal, and proportion of 15–19 year-old persons who had not finished school. Factors significantly associated with HIV infection in young men were proportion of residents who were African Blacks and proportion of an area urban, but the proportion of 20–64 year-old unemployed men in an area was marginally significant (Table 1).

In multiple variable models (Table 2), the proportion of black Africans was the dominant explanatory variable that displaced all other variables from the models, except for the proportion of an area that was urban, which remained significant in the model for men. The proportion who were black Africans and proportion who were urban were therefore retained in the model used for predicting HIV prevalence for all pixel locations for men, and the model containing only the proportion who were black Africans was used for spatial prediction of HIV prevalence in women.

Figure 2 shows estimated HIV prevalence for each pixel location for South Africa and the associated 95% confidence intervals for women and for men separately using these two models, respectively. The maps show that there is considerable variation in HIV prevalence within provinces. The lowest levels were found in inland rural areas of the Western Cape,

and the highest levels were evident in north western parts of KwaZulu Natal, southern Mpumalanga, and eastern Free State.

DISCUSSION

Disease maps have long been recognized for their potential to "provide a rapid visual summary of complex geographic information (that) may identify subtle patterns in the data,"[14] to better understand disease etiology, and to help target prevention programs. The possibilities for mapping and spatial modeling have been greatly enhanced in recent years because of advances in geographic information systems and spatial statistical modeling, and the availability of high resolution geo-referenced health data, including household surveys. Disease maps portraying the geographic distribution of chronic, non-infectious diseases, e.g., cancer atlases,[15–17] are generally based on data obtained from routine reporting systems. In the spatial epidemiology of infectious diseases such as malaria, maps have been derived from routine data where they are available,[18] but more commonly from sample surveys.[19,20] Sexually transmitted infections have previously been mapped using health service registration data.[21,22]

To our knowledge, our maps are the first HIV prevalence maps showing variation in infection levels without restriction to administrative boundaries such as provinces. On the basis of a large sample of persons, our maps depict HIV prevalence with a known estimate of uncertainty. As one might expect, prevalence of infection does not abruptly change at provincial boundaries, but instead follows a continuum across these boundaries. This should form the basis for a more integrated response to HIV prevention and treatment, rather than one based essentially on provincial confines. For example, the high prevalence cluster in the area along the common provincial boundaries between Mpumalanga and KwaZulu Natal may benefit from between province coordination of antiretroviral treatment provisions and HIV prevention campaigns.

South Africa can be divided into the high prevalence eastern half and the low prevalence western half. For women, the highest rates were evident in western parts of KwaZulu Natal near Lesotho and Swaziland, parts of southern Mpumalanga and in outlying areas of Durban. For men, the highest prevalence was in urban and in coastal areas of KwaZulu Natal, but at significantly lower levels than for women. The reasons why areas of peak prevalence in men and women of this age group do not coincide are not clear. It may be because the male partners of women in the survey were an average of four years older and were therefore often outside the age range of this survey, or it may be that many of their partners are migrant workers who work in areas with better employment prospects.[5]

Overall, young women are at approximately three times the risk for HIV infection than young men, as has been previously reported.[2,5] This difference is vividly illustrated by the two prevalence maps (Figure 2). The contrast is particularly striking in KwaZulu Natal where our model estimates a prevalence of 3% to < 5% for men compared with 25% to < 35% for women in areas of peak prevalence for women. There are many factors that place young women at higher risk of HIV infection than men; these have been identified in the main findings of this survey,[5] including the finding that young women were more likely to have older partners than young men. This would expose them more directly to the already higher levels of HIV prevalence of older cohorts in the communities from which their partners are drawn, while young men are exposed to the levels of HIV prevalence in their own cohort, which is largely HIV negative.

Our maps concur with antenatal surveillance showing that prevalence is highest in KwaZulu Natal and lowest in Western Cape and Limpopo provinces.[1] However, our maps show that

there is considerable variation within provinces: e.g., the inland rural areas of the Western Cape showed a low prevalence of < 3% in men and < 5% in women, but the risk of HIV infection in and around urban areas of Cape Town was higher at > 3–5% for men and > 8–15% for women.

The prevalence of HIV in central parts of Limpopo Province was low compared with other northeastern parts of the country. We were unable to identify any factors that would explain this relatively low prevalence, other than a very low proportion of urbanization for this province.[23]

An area being urban was independently associated with risk of HIV infection in young men after adjusting for proportion who were black Africans. An association between urban residence status and HIV infection is likely to be the result of increased mobility and possibly of a different youth culture that is less affected by parental control. Nevertheless, the survey found no difference in mean age for first time having sex between rural and urban men. Furthermore, we did not find an association between informal housing and HIV infection in young men, although we did find such an association in young women (Table 1). Such an association has been reported in other studies.[2] This negative finding may be the result of informal settlements not being homogenous in their socioeconomic and cultural milieu, or it may be that the classification available from the census is at insufficient spatial resolution to separate informal areas from those containing both formal and informal residences.

Of the potential area-based explanatory variables that can be obtained from the census, the proportion of the residents who are black Africans in an area was strongly related to HIV infection in both young women and young men. When HIV status was analyzed in relation to questionnaire derived variables from the survey,[5] a similar result was obtained even after adjusting for socioeconomic and sexual behavior variables that were available from the questionnaire. The strong association between HIV infection and the proportion who were black Africans in an area is likely to be the result of the historical differences in the HIV epidemic between Africans and other race groups in South Africa. It has been shown that the system of migrant labor during the apartheid era, which resulted in men spending long periods separated from their families, directly fueled the HIV epidemic in South Africa's black population.[24] Migration in South Africa is characterized by a high prevalence of temporary or circular migration[23] that has endured from entrenched patterns of labor migration in the past.[25] As a result, there is extensive mixing between many city dwellers and their rural counterparts that would have had a considerable impact on the pattern of the HIV epidemic in South Africa,[26] with a high prevalence of infection not only in urban centers, but also in many rural areas and small towns.

The proportion of the economically active population being unemployed was clearly related to HIV infection for both women and men (albeit marginally in the case of men), which confirmed that HIV in South Africa is linked to poor socioeconomic conditions and poverty. However, as with most of the other variables, this association did not persist once the proportion of area having black Africans was included in the model because of the correlation between unemployment and race.

It is noteworthy that the odds ratio for urban status of an area relative to a rural area and the odds ratio per percent increase in the racial composition of an area having black Africans both increased for men when adjusting for each other (Tables 1 and 2). This implies that in the single variable models, urban residence was negatively confounded by the proportion who were black Africans and *vice versa*. This is plausible because rural areas generally have more black Africans in their population than urban areas. Conversely, for young women, the

predominant risk factor for HIV infection was the racial composition of their area; living in an urban or rural area had no further effect.

A constraint to our model-based mapping approach is the limitation imposed by the restricted number of variables that are available from the national population census. To model and predict prevalence based on relevant social, demographic, and economic indicators, we require these to be available at all map locations, which generally means that only census-derived variables can be used. For instance, we would have liked to use a variable that quantifies the proportion of adults who are migrant workers, had this been available.

Other limitations of our study are those that apply to the survey from which the data are drawn, in particular, the low response rates among white youths, which may have affected our estimates in some areas. A further limitation is the computationally intense nature of the Bayesian MCMC–based modeling process, which limited the spatial resolution of the grid for which we predicted prevalence, resulting in the fairly large grid-blocks (pixels) that characterize the prediction maps.

Maps such as the ones we have produced here are based on random samples of the population and are therefore a far more reliable guide to the variation of HIV prevalence in the population than reports of much higher prevalence, which are based on volunteers who are screened for cohort studies.

The confidence intervals that accompany our map predictions will provide some assurance against over-interpretation of the data. The confidence interval maps indicate that the high prevalence among women in northwestern KwaZulu Natal may be no higher than 15%, but could be as high as 65%. Similarly, the low prevalence of HIV among women in the Western Cape, may be as high as 15%. To avoid misinterpretation of HIV prevalence maps, we would urge that these should always be accompanied by corresponding uncertainty or confidence interval maps.

The value of our maps is that they can be used as a health management tool that shows which parts of provinces are in need of particular attention. In these areas of high prevalence, further focused research should be done to inform intensified prevention and treatment programs. Because young people enter the age group largely uninfected, prevalence in this age group is a reasonable proxy for age-specific incidence of infection. Maps showing change in prevalence based on two or more surveys would provide valuable clues as to where prevention efforts are bearing fruit, and where they must be strengthened.

APPENDIX

Based on the formulation of the Bayesian geostatistical model as described by, for example, Diggle and others[27,28] and Gemperli and others,[20,29] our model is specified as follows.

Let Y_{ji} represent the binary response corresponding to the infection status of individual j at site i (the sampled EAs) taking value 1 if the youth tested HIV positive and 0 otherwise. The Y_{ji} are conditionally independent Bernoulli variables with infection probability p_i at location i. The p_i are defined via a generalized linear mixed model to take account of spatial dependence $\text{logit}(p_i) = X_i \beta + S(l_i)$ where β represents the regression coefficients for a set of known covariates X at the centroids l_i of all sampled EAs in the study area. $S = (S(l_1), \ldots, S(l_n))^T$ denotes the values of the (unobserved) Gaussian spatial process $S(\cdot)$ at sample locations l_i; $\sigma^2 = \text{Var}\{S(l)\}$, and ϕ is a parameter of the correlation function $\rho(d_{ij}, \phi)$, in our case $\exp(-d_{ij}/\phi)$, where d_{ij} is the distance between locations l_i and l_j.

Prior distributions were specified for each model parameter. For and ^{2}, flat and uniform priors were specified, respectively. For , a discrete exponential prior with mean of 0.8 and 100 discretization points in the interval 0.2–1.2 was specified.

Let S* represent the values of S(·) at other locations of interest, i.e., the grid of pixels covering the entire map. To produce a smooth map of HIV prevalence, we estimated $Y_O = (Y(l_{O1}), Y(l_{O2}), \ldots \ldots Y(l_{ON}))$ at these unsampled locations $l_O = (l_{O1}, l_{O2}, \ldots \ldots l_{ON})$ by obtaining MCMC estimates from the following predictive distribution:

$$P\left(Y_O | Y\right) = \int P\left(Y_O | \beta, S^*\right) P\left(S^* | S, \sigma^2, \Phi\right) P\left(\beta, S, \sigma^2, \Phi | Y\right) \, \mathrm{d}\beta \mathrm{d}S^* \mathrm{d}S \mathrm{d}\sigma^2 \mathrm{d}\Phi$$

where

$$P\left(Y_O | \beta, S^*\right) \sim Bn\left(p\left(\ell_{Oi}\right)\right)$$

with

$$\mathrm{logit}\left(p\left(\ell_{Oi}\right)\right) = x\left(\ell_{Oi}\right)\beta + S\left(\ell_{Oi}\right); P\left(S^* | S, \sigma^2, \Phi\right)$$

is the distribution of the spatial effects at new sites (kriging); and P(, S, ^{2}, | Y) is the posterior distribution of the model parameters.

Convergence was assessed by inspecting plots of traces of simulations for individual parameters. The first 10,000 iterations were discarded; thereafter every 50th sample was retained until 5,000 iterations had been attained, i.e. simulations were run for 250,000 iterations.

At each prediction location, the median and 2.5th percentiles and 97.5th percentiles were calculated from the MCMC simulations, and mapped, using ArcView.[30] Approximate *P* values were derived from z-scores that were calculated as the mean divided by a quarter of the 2.5th–97.5th percentile range.

REFERENCES

1. National Department of Health (South Africa). National HIV and Syphilis Antenatal Sero-Prevalence Survey in South Africa 2004. Pretoria: 2004. Available from www.health.gov.za

2. Shisana, O.; Rehle, T.; Simbayi, LC.; Parker, W.; Zuma, K.; Bhana, A.; Connolly, C.; Jooste, S.; Pillay, V. South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey. Human Sciences Research Council Press; Cape Town, South Africa: 2005.

3. Joint United Nations Programme on HIV/AIDS. Report on the Global AIDS Epidemic. UNAIDS; Geneva: 2004.

4. Erb-Leoncavallo, A.; Holmes, G.; Jacobs, G.; Urdang, S.; Vanek, J.; Zarb, M. Women and HIV/AIDS: Confronting the Crisis. The Joint United Nations Programme on HIV/AIDS; The United Nations Population Fund, The United Nations Development Fund for Women; Geneva: New York: 2004. 2004

5. Pettifor AE, Rees HV, Kleinschmidt I, Steffenson A, MacPhail C, Hlongwa-Madikizela L, Vermaak K, Padian NS. Young people's sexual health in South Africa: HIV prevalence and sexual behaviours from a nationally representative household survey. AIDS. 2005; 19:1525–1534. [PubMed: 16135907]

6. Goldstein S, Usdin S, Scheepers E, Japhet G. Communicating HIV and AIDS, what works? A report on the impact evaluation of Soul City's fourth series. J Health Commun. 2005; 10:465–483. [PubMed: 16199389]

7. Pettifor A, Kleinschmidt I, Levin J, Rees H, MacPhail C, Hlongwa-Madikizela L, Vermaak K, Napier G, Stevens W, Padian N. A community-based study to examine the effect of a youth prevention intervention on young people aged 15–24 in South Africa: results of the baseline survey. Trop Med Int Health. 2005; 10:971–980. [PubMed: 16185231]

8. Statistics South Africa. Community Profile Databases: Census 2001. 2001. Product Reference No. 03-02-22. Available from http://www.statssa.gov.za

9. Gallo D, George J, Fitchen J, Goldstein A, Hindahl M. Evaluation of a system using oral mucosal transudate for HIV-1 antibody screening and confirmatory testing. OraSure HIV Clinical Trials Group. JAMA. 1997; 277:254–258. [PubMed: 9005276]

10. Pettifor, AE.; Rees, HV.; Steffenson, A.; Hlongwa-Madikizela, L.; MacPhail, C.; Vermaak, K.; Kleinschmidt, I. HIV and Sexual Behaviour Among Young South Africans: A National Survey of 15–24 Year Olds. RHRU and LoveLife; Johannesburg: 2004.

11. Longley, P.; Goodchild, MF.; Maguire, DJ.; Rhind, DW. Geographic Information Systems and Science. John Wiley; Chichester, United Kingdom: 2001.

12. Statistics South Africa. Census 2001 Products: Small Area Statistics. 2001. Reference number: 03-02-45. Available from http://www.statssa.gov.za/census01/html/C2001smallareastats.asp

13. Christensen OF, Ribeiro PJ Jr. GeoRglm: a package for generalised linear spatial models. R News. 2002; 2:26–28.

14. Elliott P, Wartenberg D. Spatial epidemiology: current approaches and future challenges. Environ Health Perspect. 2004; 112:998–1006. [PubMed: 15198920]

15. Swerdlow, A.; dos Santos Silva, I. Atlas of Cancer Incidence in England and Wales 1968–85. Oxford University Press; Oxford, United Kingdom: 1993. 1993

16. Devesa, SS.; Grauman, DG.; Blot, WJ.; Pennello, G.; Hoover, RN.; Fraumeni, JF, Jr. Atlas of Cancer Mortality in the United States, 1950–1994. U.S. Government Printing Office; Washington, DC: 1999.

17. Pickle, LW.; Mungiole, M.; Jones, GK.; White, AA. Atlas of United States Mortality. National Center for Health Statistics; Hyattsville, MD: 1996.

18. Kleinschmidt I, Sharp B, Mueller I, Vounatsou P. Rise in malaria incidence rates in South Africa: a small area spatial analysis of variation in time trends. Am J Epidemiol. 2002; 155:257–264. [PubMed: 11821251]

19. Kleinschmidt I, Omumbo J, Briët O, Van de Giesen N, Sogoba N, Mensah NK, Windmeijer P, Moussa M, Teuscher T. An empirical malaria distribution map for west Africa. Trop Med Int Health. 2001; 6:779–786. [PubMed: 11679126]

20. Gemperli A, Vounatsou P, Kleinschmidt I, Bagayoko M, Lengeler C, Smith T. Spatial patterns of infant mortality in Mali: the effect of malaria endemicity. Am J Epidemiol. 2004; 159:64–72. [PubMed: 14693661]

21. Kleinschmidt I, Ramkissoon A, Morris N, Mabude Z, Curtis B, Beksinska M. Mapping indicators of sexually transmitted infection services in the South African public health sector. Trop Med Int Health. 2006; 11:1047–1057. [PubMed: 16827705]

22. Law DCG, Serre ML, Christakos G, Leone PA, Miller WC. Spatial analysis and mapping of sexually transmitted diseases to optimise intervention and prevention strategies. Sex Transm Infect. 2004; 80:294–299. [PubMed: 15295129]

23. Kok, P.; Collinson, M. Migration and Urbanization in South Africa. Statistics South Africa; Pretoria: 2006. Report 03-04-02

24. Lurie M, Williams B, Zuma K, Mkaya-Mwamburi D, Garnett G, Sturm AW, Sweat M, Gittelsohn J, Abdool Karim S. The impact of migration on HIV-1 transmission: a study of migrant and non-migrant men, and their partners. Sex Transm Dis. 2003; 30:149–156. [PubMed: 12567174]

25. Collinson, M.; Kok, P.; Garenne, M. Migration and Changing Settlement Patterns: Multilevel Data for Policy. Statistics South Africa; Pretoria: 2006. Report 03-04-01

26. Lurie MN, Williams BG, Zuma K, Mkaya-Mwamburi D, Garnett GP, Sweat MD, Gittelsohn J, Karim SS. Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa. AIDS. 2003; 17:2245–2252. [PubMed: 14523282]

27. Diggle PJ, Tawn JA, Moyeed R. Model-based geostatistics. J R Stat Soc C. 1998; 47:299–350.

28. Diggle, PJ.; Ribeiro, PJ.; Christensen, OF. An introduction to model-based geostatistics. In: Moller, J., editor. Spatial Statistics and Computational Methods. Springer Verlag; New York: 2003. p. 43-86.

29. Gemperli A, Vounatsou P. Fitting generalised linear mixed models for point-referenced spatial data. J Modern Appl Stat Methods. 2003; 2:497–511.

30. Environmental Sciences Research Institute. ArcMap™. Version 9.0. Environmental Sciences Research Institute; Redlands, CA: 2004.
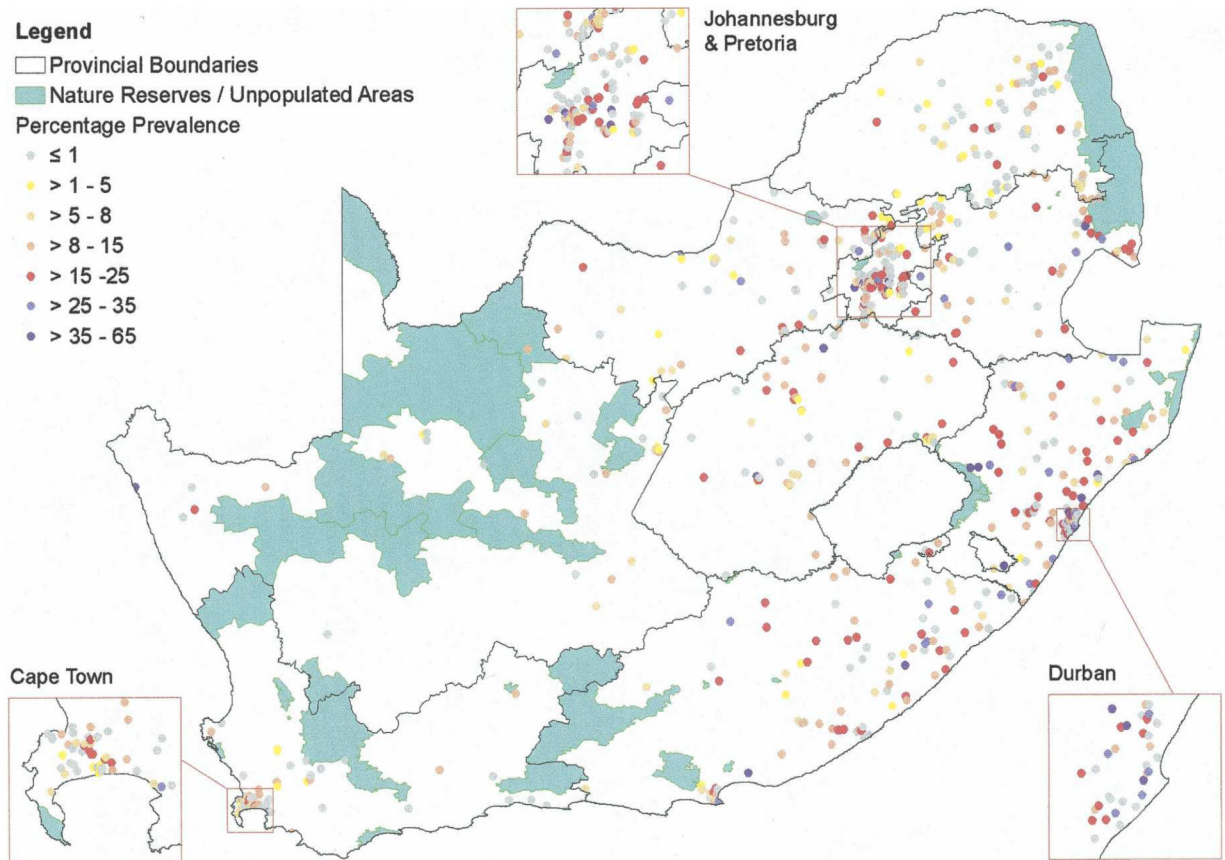
Figure 1.
Observed prevalence of human immunodeficiency virus in 15–24 year-old persons at survey
sample locations in South Africa, 2003.
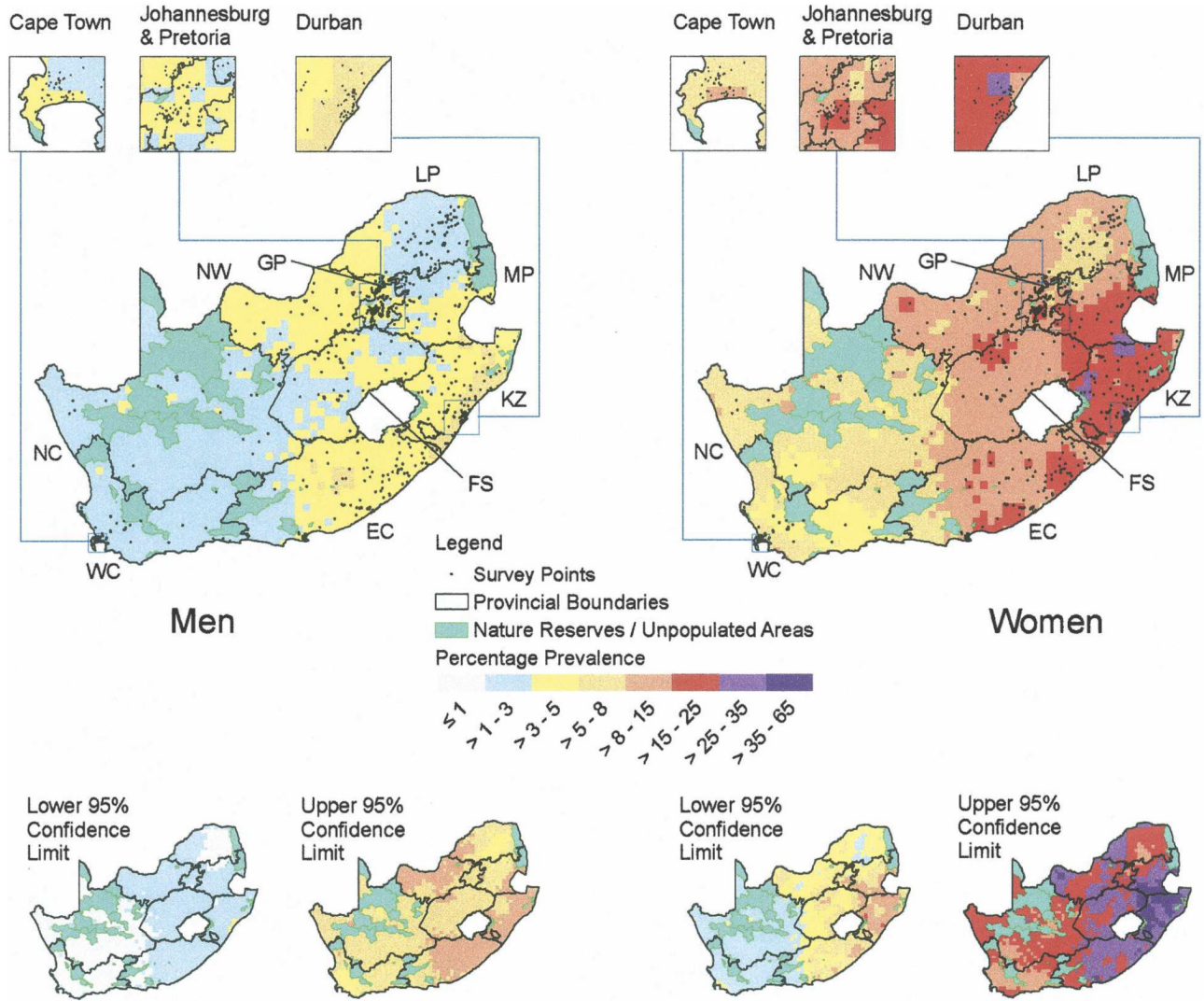
Figure 2.
Prevalence of human immunodeficiency virus and 95% confidence limits predicted from spatial models for 15–24 year-old persons in South Africa, 2003. LP = Limpopo Province; MP = Mpumalanga; KZ = KwaZulu Natal; FS = Freestate; EC = Eastern Cape; WC = Western cape; NC = Northern Cape; NW = North West; GP = Gauteng Province.

Table 1

Univariate association between HIV infection and area characteristic; Parameter estimates from single variable logistic regression spatial models of HIV prevalence in 15–24 year-old persons in South Africa, by sex [*]

Area characteristic	Odds ratios		*P*
	Median	**95% CI**	
Women			
Urban residence, relative to rural	0.94	0.75–1.22	0.65
Proportion African Black, per %	1.011	1.007–1.014	< 0.0001
Proportion of 20–64 year-old persons unemployed, per %	1.0062	1.0007–1.0117	0.032
Population density, per 1,000 per km^2	1.0002	0.9999–1.0004	0.24
Proportion of households informal, per %	1.0037	1.0011–1.0063	0.006
Proportion of 15–19 year-old persons dropped out of school, per %	1.0051	1.0005–1.0097	0.031
Men			
Urban residence, relative to rural	1.59	1.13–2.21	0.007
Proportion African Black, per %	1.0048	1.0010–1.0099	0.028
Proportion of 20–64 year-old persons unemployed, per %	1.0083	1.0001–1.0170	0.062
Population density, per 1,000 per km^2	1.0002	0.9999–1.0005	0.21
Proportion of households informal, per %	1.0005	0.9957–1.0045	0.87
Proportion of 15–19 year-old persons dropped out of school, per %	1.0045	0.9980–1.0128	0.20

[*]
HIV = human immunodeficiency virus; CI = confidence interval.

Table 2

Association between HIV infection and area characteristic; Parameter estimates from logistic spatial regression models of HIV prevalence in 15–24 year-old persons used for spatial prediction[*]

Area characteristic	Model coefficients		Odds ratios		
	Median	95% CI	Median	95% CI	*P*
Women					
Proportion African Black, per %	0.0104	0.0070–0.0136	1.0105[†]	1.0071–1.0137	< 0.0001
	0.75	0.40–1.16			
	0.353	0.213–0.578			
Men					
Urban residence, relative to rural, per %	0.627	0.253–0.950	1.8715	1.2884–2.5857	0.0005
Proportion African Black	0.0073	0.0022–0.0117	1.0073[‡]	1.0022–1.0118	0.0016
	0.98	0.40–1.19			
2	0.222	0.086–0.519			

[*] HIV = human immunodeficiency virus; CI = confidence interval.

[†] Equivalent to an odds ratio of 1.23 per 20% increase in proportion Black.

[‡] Equivalent to an odds ratio of 1.16 per 20% increase in proportion Black.