



Published in final edited form as:

AIDS Care. 2016 February ; 28(2): 266–272. doi:10.1080/09540121.2015.1080793.

A county-level analysis of persons living with HIV in the southern United States

Simone C. Gray^a, Tyler Massaro^b, Isabel Chen^c, Christina J. Edholm^d, Rachel Grotheer^e, Yiqiang Zheng^{f,g}, and Howard H. Chang^h

^aDivision of HIV/AIDS and Prevention, Centers for Disease Control and Prevention, Atlanta, GA, USA

^bDepartment of Mathematics, The University of Tennessee, Knoxville, TN, USA

^cDepartment of Mathematics, Emory University, Atlanta, GA, USA

^dDepartment of Mathematics, University of Nebraska-Lincoln, Lincoln, NE, USA

^eDepartment of Mathematical Sciences, Clemson University, Clemson, SC, USA

^fDepartment of Mathematics, Purdue University, West Lafayette, IN, USA

^gDepartment of Statistics, Purdue University, West Lafayette, IN, USA

^hDepartment of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA

Abstract

This study uses county-level surveillance data to systematically analyze geographic variation and clustering of persons living with diagnosed HIV (PLWH) in the southern United States in 2011. Clusters corresponding to large metropolitan areas – including Miami, Atlanta, and Baltimore – had HIV prevalence rates higher ($p < .001$) than the regional rate. Regression analysis within the counties included in these clusters determined that race was a significant indicator for PLWH. These results provide a general picture of the distribution of PLWH in the southern United States at the county level and provide insights for identifying local geographic areas with a high number of PLWH, as well as subpopulations that may have an increased risk of infection.

Keywords

HIV; race/ethnicity; clusters; socioeconomic status; American South

Introduction

Rates of HIV/AIDS differ by population and geographic region in the United States (US). The disproportionate impact of HIV/AIDS on the southern US and in disadvantaged groups

CONTACT Simone C. Gray ✉ simonegray@cdc.gov.

Disclosure statement

No potential conflict of interest was reported by the authors.

has been well documented (O’Leary, Broadwell, Yao, & Hasin, 2006; Prejean, Tang, & Hall, 2013; Reif et al., 2014, 2015; Smith, 2006; Thomas, 2006). Trends in HIV have demonstrated a higher prevalence in blacks/African-Americans (blacks), and among persons with low income and low education attainment (Gant et al., 2012; Reif, Golin, & Smith, 2005; Song et al., 2011). With an over-representation of these subpopulations in the South, marked disparities in HIV prevalence exist (Reif, Geonnotti, & Whetten, 2006; Thomas, 2006). In 2011, the estimated rate of diagnoses of HIV infection among blacks was nearly 8 times as high as the rate for non-Hispanic whites (whites) in the southern US (Centers for Disease Control and Prevention, 2013).

To date, most studies of HIV infection and at-risk populations have used aggregated population data at the national, regional, or state level (Aral, O’Leary, & Baker, 2006; Hanna, Selik, Tang, & Gange, 2012; Zeglin & Stein, 2015). More recent studies on HIV/AIDS and social determinants of health (SDH) have been conducted on a finer scale, including zip code-, county-, and community-level as data are becoming more accessible at finer resolutions (Harrison, Ling, Song, & Hall, 2008; Rubin, Colen, & Link, 2009; Trepka et al., 2013). Results at this level are potentially more useful to local healthcare professionals, while offering relevant information for enacting programs related to HIV prevention and care (Song et al., 2011). This study provides a general picture of the distribution of persons living with diagnosed HIV (PLWH) in the southern US using county-level HIV surveillance and U.S. Census data.

Methods

HIV surveillance and demographic data

County-level estimates of prevalence rates (per 100,000), as well as the number of PLWH in 2011 were obtained from the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention Atlas at the U.S. Centers for Disease Control and Prevention (CDC) (NCHHSTP Atlas, 2014). The analysis was based on 16 states in the US south: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

To ensure confidentiality of cases, county-level totals were suppressed for counties with populations of less than 100 or fewer than 5 PLWH and in accordance with approved state requirements under data agreements with CDC (NCHHSTP Atlas, 2014). To gain an overall picture of PLWH in the south and facilitate spatial analysis, we imputed the missing data for each data-suppressed county. To impute missing prevalence rates, we calculated state-specific average prevalence rates using the total number of reported cases and the total population size among reporting counties (U.S. Census Bureau American Community Survey: 2008–2012 5-Year Estimates, 2014).

We obtained data for the demographic variables from the 2008–2012 American Community Survey 5-Year Estimates (U.S. Census Bureau American Community Survey: 2008–2012 5-Year Estimates, 2014). County-level SDH variables obtained included poverty status (percent of county population below the federal poverty level), educational attainment (percent of county population with less than high school education), income (median

household income), urban indicator (assigned counties a value of 1 if they had at least 50,000 inhabitants, and 0 otherwise), unemployment rate, percent black, percent white, and percent Hispanic (Census, 2013).

Statistical analysis

We used global and local clustering analyses to detect overall and specific clusters of elevated counts of PLWH, respectively. Two measures of global spatial dependence, Moran's I index (Moran, 1950) and Geary's contiguity ratio c (Geary, 1954) were used to provide an indication of the presence of spatial clustering. The Besag and Newell cluster (BNC) detection (Besag & Newell, 1991) was used to identify the locations and extent of clusters. The BNC method considers windows with a predetermined number of cases, k , representing the size of the cluster. The analysis was repeated for different values of k to detect recurring clusters. For each value of k , we selected the top nine clusters defined as having a prevalence ratio greater than two and the smallest p -values (all $< .05$). We then selected the counties that appeared at least twice over different values of k .

We considered three regression approaches to examine associations between PLWH prevalence and demographic variables. First, we treated log-transformed county-specific rates as a continuous outcome in linear regression models for each covariate separately and jointly. We then extended the linear regression model to include spatially dependent random intercepts to account for potentially spatially varying unmeasured confounders using a conditional autoregressive (CAR) model where spatial dependence is based on county neighborhoods (Waller & Gotway, 2004). Lastly we modeled the cases as counts using Poisson regression, under a Bayesian hierarchical framework with CAR random effects. We assume uninformative normal priors for the coefficients and uninformative inverse gamma priors for the variances. All three approaches were performed on all counties in the US south and on the clusters identified by the BNC. All statistical analysis was conducted using R3.1.1 software (R Core Team, 2014).

Results

Summaries of PLWH prevalence

A total of 1422 counties were included with imputed data for 168 counties (11.8%). Figure 1 shows the county-level distribution of PLWH rates in the south. Counties with the highest prevalence rates were found in Texas (Walker), Florida (Union) and Maryland (Baltimore City). The prevalence rate per 100,000 for PLWH in the US south in 2011 was 307.9 with Florida having the highest rate of 446.9 (Table 1). Florida also had the second highest percent of Hispanics and the highest unemployment rate. West Virginia had one of the lowest PLWH rates (72.0 per 100,000) and the lowest percent of blacks and Hispanics.

Clustering analysis

Results of the global clustering indicate positive spatial dependence in the rates of PLWH: Moran's index $I = 0.258$ ($p < .001$) and Geary's contiguity ratio $c = 0.632$ ($p = .07$). The BNC analysis identified 78 counties with high prevalence of PLWH for multiple values of $k = 50, 100, 300, 500, 1000, 15,000, 20,000, \text{ and } 25,000$ (Figure 2). The clusters capture select

regions of Tennessee, Alabama, Georgia, Florida, Maryland, Virginia, and Delaware. Using nonparametric Wilcoxon rank sum tests, all of the race and SDH indicators, with the exception of the unemployment rates, in the clusters identified by BNC, were significantly different from the regional measures ($p < .05$). A Wilcoxon rank sum test also indicated that the prevalence rate of PLWH in the BNC clusters (846.8 per 100,000) was significantly different ($p < .001$) from the regional prevalence rate of 307.9 per 100,000.

Regression analysis

Table 2 gives the estimated risk ratios (RR) and 95% confidence intervals for the univariate associations between SDH factors and PLWH rates obtained by standard linear regression and normal CAR models, as well as the associations between SDH factors and PLWH cases in Poisson CAR models. All models indicate that percent black was significant and positively associated with PLWH ($p < .05$). Poverty status, urbanicity, and unemployment were statistically significant and positively associated with the prevalence of PLWH in all the regression models. Income was not statistically significant in the standard linear regression model; however, when spatial effects were considered, it was negatively associated with PLWH.

Educational attainment was statistically significant and negatively associated with PLWH in the linear regression model, not statistically significant in the normal CAR model, and significant and positively associated with PLWH in the Poisson CAR model. Two factors may have contributed to the differences in association estimates between the normal and Poisson CAR model. First, the two models make different distributional assumptions on the observed log rates or counts. Second, while the coefficients obtained from both models can be interpreted as RR, the normal model estimates the difference in expected log rates and the Poisson model estimates the log of the difference in rates.

Many of the SDH factors were significantly correlated with each other. Educational attainment was significantly correlated with income; unemployment; poverty; and percent white, black and Hispanic ($p < .05$). To address multicollinearity issues, we retained only variables that were not significantly correlated with each other in the multivariate model: percent black and the urban indicator. Table 3 reports the results for the linear, normal CAR, and Poisson CAR models. Percent black and the urban indicator were both statistically significant and positively related to PLWH for the US south in all models. We performed multiple linear regression analysis on the spatial clusters identified by BNC. The results for percent black are similar to those for the entire region (Table 3).

Conclusion

This analysis is the first county-level analysis examining SDH, race, and PLWH in the southern US. Overall, the BNC analysis results showed that counties in large metropolitan areas, including Memphis, Miami, Atlanta, and Baltimore had higher rates of PLWH. From the regression analysis, we identified several SDH factors associated with PLWH rates and cases, after controlling for potential residual spatial confounding. Particularly, the positive association between county-level black proportions and PLWH rates persists even within clusters of high HIV prevalence.

There were limitations with the data and analysis in this study. We used the demographic variables restricted to the adult population of 18 and older, since the 13–18-year-old population constitutes a small percentage of each county. This restriction gives a more accurate description of PLWH. One of the problems we encountered while performing spatial clustering analysis with the BNC method was deciding on the window size, k . The subjectivity associated with choosing a value of k to use is a problem that has been discussed in previous literature (Costa & Assunção, 2005; Tango, 2010). We included multiple values of k to detect recurring clusters as suggested by Costa and Assunção (2005).

Epidemiologic evidence shows that individual, social, and structural factors (e.g., educational attainment, unemployment, and poverty rates) have been linked to the increased mortality and morbidity of diseases, with blacks being disproportionately affected by these social and economic barriers (Aral, Adimora, & Fenton, 2008; Aral et al., 2006; Davids, Hutchins, Jones, & Hood, 2014; Jones, 2001; Murray et al., 2006). Consistent with our results, previous studies at the state level have shown that racial and social disparities in the south exist in HIV incidence and PLWH (Farley, 2006; Prejean et al., 2013; Reif et al., 2006, 2015).

This analysis provides a general picture of the distribution of PLWH in the south. County-level data can be used to better inform CDC's HIV prevention approach to focus on targeting populations in specific geographic areas (Hall et al., 2015). Examining the rates of PLWH and mapping counties with higher rates and greater disparities is an important public health strategy that will allow public health professionals to (1) target the areas and populations most in need of HIV-related care, (2) allocate and distribute resources to communities most in need, and (3) prevent further transmission of the disease.

Acknowledgements

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Funding

This material was based upon work partially supported by the National Science Foundation under Grant DMS-1127914 to the Statistical and Applied Mathematical Sciences Institute.

References

- Aral SO, Adimora AA, Fenton KA. Understanding and responding to disparities in HIV and other sexually transmitted infections in African Americans. *The Lancet*. 2008; 372:337–340.
- Aral SO, O'Leary A, Baker C. Sexually transmitted infections and HIV in the Southern United States: An Overview. *Sexually Transmitted Diseases*. 2006; 33(7):S1–S5. [PubMed: 16794550]
- Besag J, Newell J. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society*. 1991; 154(1):143–155.
- Census, U. S. Geography: Urban and rural classification. 2013. Retrieved from <https://www.census.gov/geo/reference/urban-rural.html>
- Centers for Disease Control and Prevention. HIV surveillance report 2011. 2013. Retrieved from <http://www.cdc.gov/hiv/topics/surveillance/resources/reports/>

- Costa MA, Assunção RM. A fair comparison between the spatial scan and the Besag-Newell disease clustering tests. *Environmental and Ecological Statistics*. 2005; 12:301–319.
- Davids B-O, Hutchins SS, Jones CP, Hood JR. Disparities in life expectancy across US counties linked to county social factors, 2009 community health status indicators (CHSI). *Journal of Racial and Ethnic Health Disparities*. 2014; 1(1):2–11.
- Farley TA. Sexually transmitted diseases in the Southeastern United States: Location, race, and social context. *Sexually Transmitted Diseases*. 2006; 33(Supplement):S58–S64. [PubMed: 16432486]
- Gant Z, Lomotey M, Hall HI, Hu X, Guo X, Song R. A county-level examination of the relationship between HIV and social determinants of health: 40 states, 2006–2008. *AIDS Journal*. 2012; 6:1–7.
- Geary RC. The contiguity ratio and statistical mapping. *The Incorporated Statistician*. 1954; 5(3):115–146.
- Hall I, An Q, Tang T, Song R, Chen M, Green T, Kang J. Prevalence of diagnosed and undiagnosed HIV infection - United States, 2008–2012. *MMWR Morbidity and Mortality Weekly Report*. 2015; 64:657–662. [PubMed: 26110835]
- Hanna DB, Selik RM, Tang T, Gange SJ. Disparities among states in HIV-related mortality in persons with HIV infection, 37 U.S. States, 2001–2007. *AIDS*. 2012; 26:95–103. [PubMed: 22008659]
- Harrison KM, Ling Q, Song R, Hall HI. County-level socioeconomic status and survival after HIV diagnosis, United States. *Annals of Epidemiology*. 2008; 18(12):919–927. [PubMed: 19041591]
- Jones CP. ‘Race’, racism and the practice of epidemiology. *American Journal of Epidemiology*. 2001; 154(4):299–304. [PubMed: 11495851]
- Moran PAP. Notes on continuous stochastic phenomena. *Biometrika*. 1950; 37(1):17–23. [PubMed: 15420245]
- Murray CJL, Kulkarni SC, Michaud C, Tomijima N, Bulzacchelli MT, Iandiorio TJ, Ezzati M. Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine*. 2006; 3(9):1513–1524.
- NCHHSTP Atlas. 2014. Retrieved July 2014, from <http://www.cdc.gov/nchhstp/atlas/>
- O’Leary A, Broadwell SD, Yao P, Hasin D. Major depression, alcohol and drug use disorders do not appear to account for the sexually transmitted disease and HIV epidemic in the Southern United States. *Sexually Transmitted Diseases*. 2006; 33(7):S70–S77. [PubMed: 16543865]
- Prejean J, Tang T, Hall HI. HIV diagnoses and prevalence in the southern region of the United States, 2007–2010. *Journal of Community Health*. 2013; 38:414–426. [PubMed: 23179388]
- R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. Retrieved from <http://www.R-project.org/>
- Reif S, Geonotti KL, Whetten K. HIV infection and AIDS in the Deep South. *American Journal of Public Health*. 2006; 96(6):970–973. [PubMed: 16670228]
- Reif S, Golin CE, Smith SR. Barriers to accessing HIV/AIDS care in North Carolina: Rural and urban differences. *AIDS Care*. 2005; 17(5):558–565. [PubMed: 16036242]
- Reif S, Pence BW, Hall I, Hu X, Whetten K, Wilson E. HIV diagnoses, prevalence and outcomes in nine southern states. *Journal of Community Health*. 2015; 40:642–651. [PubMed: 25524210]
- Reif SS, Whetten K, Wilson ER, McAllaster C, Pence BW, Legrand S, Gong W. HIV/AIDS in the Southern USA: A disproportionate epidemic. *AIDS Care*. 2014; 26:351–359. [PubMed: 23944833]
- Rubin MS, Colen CG, Link BG. Examination of inequalities in HIV/AIDS mortality in the United States from a fundamental cause perspective. *American Journal of Public Health*. 2009; 100:1053–1059. [PubMed: 20403885]
- Smith DB. Racial disparities in care: The concealed legacy of a divided system. *Sexually Transmitted Diseases*. 2006; 33(7):S65–S69. [PubMed: 16794557]
- Song R, Hall HI, Harrison KM, Sharpe TT, Lin LS, Dean HD. Identifying the impact of social determinants of health on disease rates using correlation analysis of area-based summary information. *Public Health Reports*. 2011; 126:70–80. [PubMed: 21836740]
- Tango, T. *Statistical methods for disease clustering*. New York: Springer science & business media; 2010.

- Thomas JC. From slavery to incarceration: social forces affecting the epidemiology of sexually transmitted diseases in the rural south. *Sexually Transmitted Diseases*. 2006; 33(7):S6–S10. [PubMed: 16794556]
- Trepka MJ, Niyonsenga T, Maddox L, Lieb S, Lutfi K, Pavlova-McCalla E. Community poverty and trends in racial/ethnic survival disparities among people diagnosed With AIDS in Florida, 1993–2004. *American Journal of Public Health*. 2013; 103:717–726. [PubMed: 23409892]
- U.S. Census Bureau American Community Survey: 2008–2012 5-Year Estimates. 2014. Retrieved from <http://factfinder2.census.gov>
- Waller, LA.; Gotway, CA. *Applied spatial statistics for public health data*. Hoboken, NJ: Wiley-Interscience; 2004.
- Zeglin RJ, Stein JP. Social determinants of health predict state incidence of HIV and AIDS: A short report. *AIDS Care*. 2015; 27:255–259. [PubMed: 25225050]

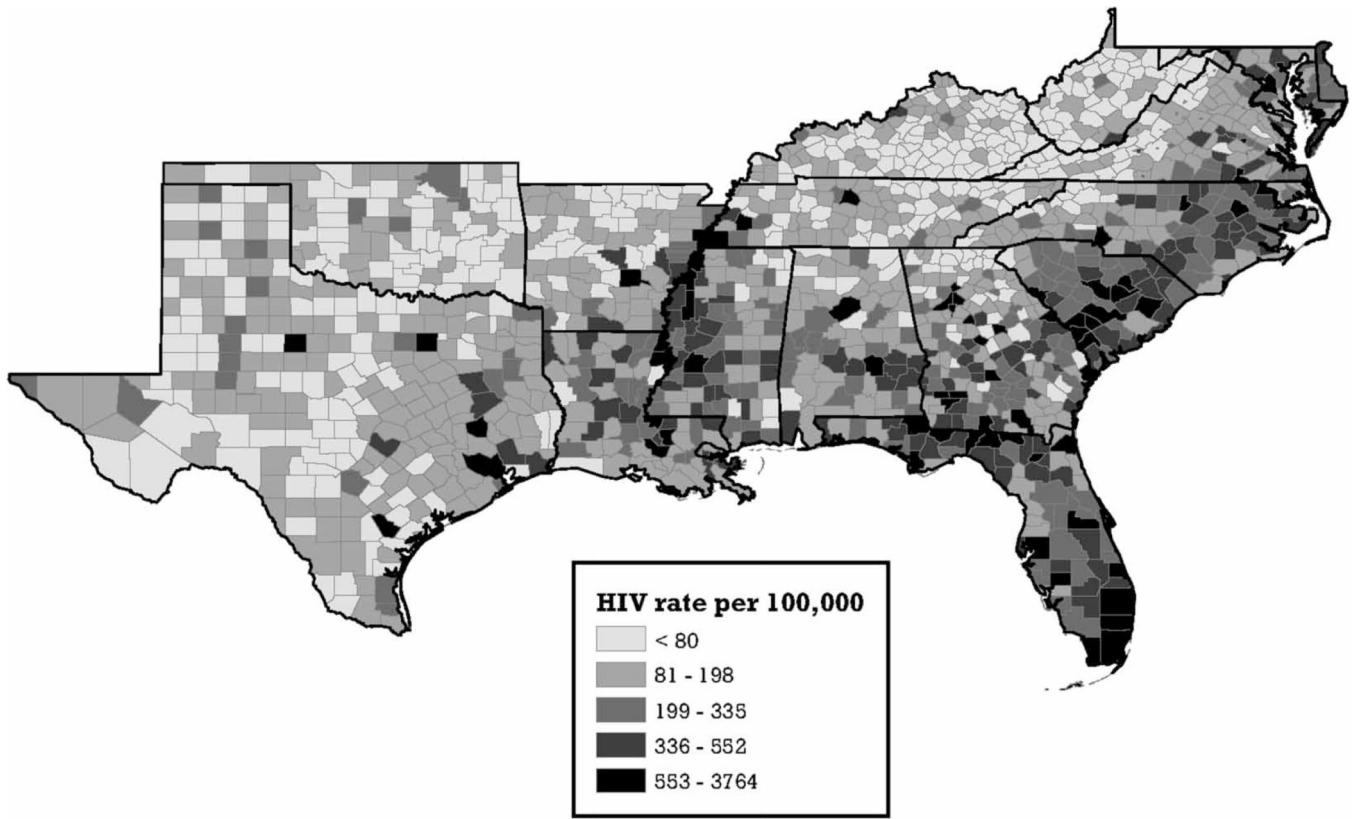


Figure 1. County-level rates per 100,000 of persons living with HIV in the U.S. South, 2011.

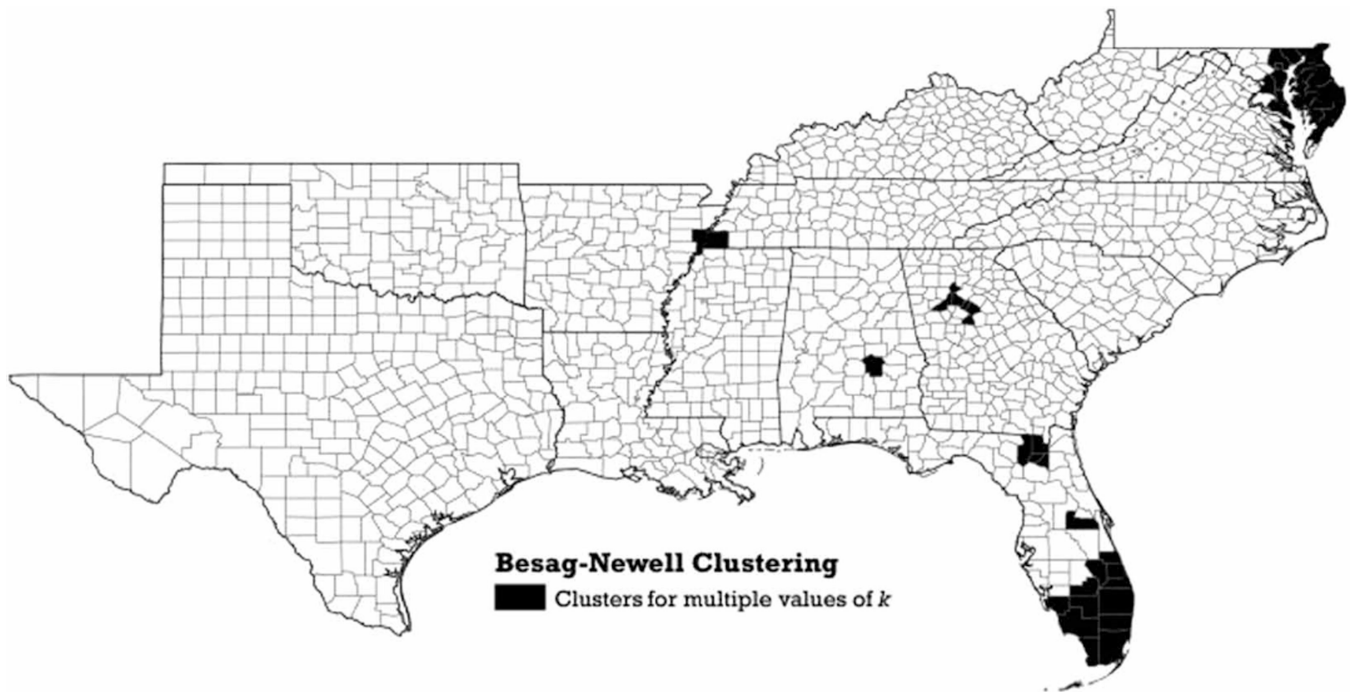


Figure 2.
Map of Besag-Newell clustering in the U.S. South.

Table 1

Descriptive statistics of 2011 PLWH rates^a per 100,000 and 2008–2012 U.S. Census data for the U.S. South.

	U.S. South	AL	AR	DE	FL	GA	KY	LA	MD	MS	NC	OK	SC	TN	TX	VA	WV
PLWH Rates																	
Mean	307.9	204.7	131.8	333.4	446.9	245.6	71.7	293.8	358.9	266.0	230.2	84.1	375.9	118.3	142.8	231.0	72.0
SD	182.7	143.6	116.7	100.3	439.2	238.6	55.4	266.6	473.5	176.7	146.5	54.4	185.3	126.1	252.5	235.8	69.9
% Black																	
Mean	18.7	26.2	15.4	20.9	15.2	30.2	7.7	31.8	29.0	37.0	21.2	7.1	27.7	16.6	11.5	19.1	3.1
SD	18.0	22.4	17.8	5.1	9.4	17.5	4.2	14.4	16.4	20.7	16.3	3.5	16.5	10.4	6.8	16.7	2.5
% White																	
Mean	60.1	67.0	74.5	65.3	57.8	55.8	86.3	60.3	54.7	58.0	65.2	68.7	64.1	75.6	45.3	64.8	93.0
SD	20.3	20.7	17.1	5.9	14.6	17.0	5.9	13.5	19.4	19.8	17.5	9.3	15.1	11.5	21.2	18.0	3.8
% Hispanic																	
Mean	15.9	3.8	6.4	8.1	22.5	8.8	3.0	4.3	8.2	2.6	8.3	8.8	5.0	4.5	37.6	7.9	1.2
SD	15.3	2.9	4.9	1.3	11.9	5.2	1.6	2.0	3.6	1.9	3.8	6.6	3.0	2.3	23.1	5.2	1.0
% Less than High School																	
Mean	16.0	17.3	16.7	12.3	14.2	15.6	17.5	17.8	11.6	18.9	15.4	13.8	15.9	16.1	19.3	13.1	16.5
SD	6.7	4.9	4.7	1.9	6.7	6.2	7.4	5.7	4.0	6.0	5.0	4.1	5.1	5.2	8.2	6.6	6.0
% Below Poverty																	
Mean	14.2	15.6	16.0	9.9	13.8	15.2	16.4	16.1	8.6	19.0	14.6	14.3	15.2	15.0	14.6	10.2	15.9
SD	5.8	5.3	3.8	0.3	4.5	5.8	6.3	5.3	4.3	6.1	4.1	4.1	4.6	4.2	5.4	6.7	4.1
Income (\$)																	
Mean	42,164	37,812	35,855	57,716	43,875	40,384	37,922	41,977	68,995	34,473	41,673	42,166	39,229	39,180	44,957	52,561	37,781
SD	12,244	8,310	5,943	5,076	7,372	11,498	10,109	9,229	19,758	7,773	7,596	7,250	7,855	8,862	10,344	19,424	6,755
% Urban																	
Mean	29.9	40.3	18.7	100.0	61.2	25.8	13.8	35.9	70.8	17.1	53.0	19.5	52.2	30.5	24.4	26.1	20.0
Employment																	
Mean	9.4	10.4	8.7	8.4	11.4	10.7	9.6	8.5	7.9	10.8	10.6	6.8	11.2	9.9	7.8	7.1	7.9
SD	3.6	4.0	2.8	0.5	2.7	3.2	2.9	3.3	2.1	4.2	2.7	2.9	2.9	2.6	2.8	2.9	2.4

Note: Data include persons with a diagnosis of HIV infection regardless of stage of disease at diagnosis.

^aRates are per 100,000.

Table 2

Estimated RR and 95% confidence intervals for the univariate associations between the SDH variables and rates of persons living with HIV (PLWH) obtained from linear and normal CAR models, and SDH factors and PLWH cases in Poisson CAR models.

	Linear	Normal CAR	Poisson CAR
Non-Hispanic Black	1.03* (1.03, 1.04)	1.03* (1.03, 1.04)	1.01* (1.01, 1.02)
Non-Hispanic White	0.98* (0.97, 0.98)	0.98* (0.97, 0.98)	0.97* (0.97, 0.98)
Hispanic	1.00 (1.00, 1.00)	1.00 (1.00, 1.01)	0.98* (0.97, 0.98)
Less than High School Education	0.99* (0.98, 1.00)	1.00 (0.99, 1.00)	1.06* (1.04, 1.08)
Below poverty level	1.02* (1.01, 1.03)	1.03* (1.02, 1.04)	1.12* (1.10, 1.13)
Med. Income (per \$10,000)	1.04 (1.00, 1.08)	0.95* (0.91, 0.99)	0.63* (0.56, 0.74)
Urban indicator	1.68* (1.51, 1.85)	1.45* (1.32, 1.59)	1.30* (1.09, 1.64)
Unemployment	1.07* (1.05, 1.08)	1.04* (1.02, 1.05)	1.04* (1.02, 1.05)

* $p < .05$.

Table 3

Estimated RR and 95% confidence intervals for the multivariate associations between SDH variables and rates of persons living with HIV (PLWH) obtained from linear and Normal CAR models, and SDH factors and PLWH cases in Poisson CAR models.

		Linear	Normal CAR	Poisson CAR
Entire US south	% Black	1.03* (1.03, 1.04)	1.03* (1.03, 1.04)	1.03* (1.03, 1.04)
	Urban indicator	1.66* (1.54, 1.79)	1.46* (1.35, 1.58)	1.35* (1.13, 1.55)
BNC	% Black	1.02* (1.01, 1.04)	1.03* (1.02, 1.03)	1.04* (1.03, 1.05)
	Urban indicator	0.94 (0.58, 1.54)	0.81 (0.64, 1.02)	1.11 (0.75, 1.39)

* $p < .05$.