

**Appendix Figure 1: Crude COVID-19 infection rates per 10,000 population as of Oct 7, 2020**



**Appendix Figure 2: Prevalent HIV infection rate in 2018 per 10,000 population**

**Appendix Table 1: Counties detected as hotspots of both COVID-19 and HIV infections (i.e., category Yellow in Figure 2)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **County** | **State** | **Rank** | **Identified as an Ending the HIV epidemic county** |
| 1 | Bullock | Alabama | 40 |  |
| 2 | Dallas | Alabama | 52 |  |
| 3 | Hale | Alabama | 66 |  |
| 4 | Lowndes | Alabama | 42 |  |
| 5 | Montgomery | Alabama | 1 |  |
| 6 | Crittenden | Arkansas | 45 |  |
| 7 | Jefferson | Arkansas | 1 |  |
| 8 | Lee | Arkansas | 56 |  |
| 9 | Baker | Florida | 49 |  |
| 10 | Broward | Florida | 1 | Yes |
| 11 | Columbia | Florida | 1 |  |
| 12 | DeSoto | Florida | 55 |  |
| 13 | Escambia | Florida | 42 |  |
| 14 | Gadsden | Florida | 1 |  |
| 15 | Glades | Florida | 68 |  |
| 16 | Hamilton | Florida | 1 |  |
| 17 | Hendry | Florida | 62 |  |
| 18 | Jackson | Florida | 1 |  |
| 19 | Jefferson | Florida | 73 |  |
| 20 | Madison | Florida | 1 |  |
| 21 | Miami-Dade | Florida | 1 | Yes |
| 22 | Suwannee | Florida | 54 |  |
| 23 | Taylor | Florida | 1 |  |
| 24 | Union | Florida | 1 |  |
| 25 | Washington | Florida | 1 |  |
| 26 | Bacon | Georgia | 72 |  |
| 27 | Bibb | Georgia | 42 |  |
| 28 | Charlton | Georgia | 74 |  |
| 29 | Clarke | Georgia | 67 |  |
| 30 | Clinch | Georgia | 60 |  |
| 31 | Colquitt | Georgia | 1 |  |
| 32 | Decatur | Georgia | 1 |  |
| 33 | Glynn | Georgia | 53 |  |
| 34 | Hancock | Georgia | 61 |  |
| 35 | Randolph | Georgia | 48 |  |
| 36 | Stewart | Georgia | 1 |  |
| 37 | Tift | Georgia | 47 |  |
| 38 | Ware | Georgia | 1 |  |
| 39 | Wheeler | Georgia | 69 |  |
| 40 | Allen | Louisiana | 1 |  |
| 41 | Avoyelles | Louisiana | 70 |  |
| 42 | Calcasieu | Louisiana | 64 |  |
| 43 | East Carroll | Louisiana | 71 |  |
| 44 | East Feliciana | Louisiana | 1 |  |
| 45 | Iberville | Louisiana | 1 |  |
| 46 | Jefferson | Louisiana | 1 |  |
| 47 | Ouachita | Louisiana | 1 |  |
| 48 | St. Landry | Louisiana | 57 |  |
| 49 | West Feliciana | Louisiana | 1 |  |
| 50 | Bolivar | Mississippi | 63 |  |
| 51 | Coahoma | Mississippi | 1 |  |
| 52 | Copiah | Mississippi | 58 |  |
| 53 | Holmes | Mississippi | 59 |  |
| 54 | Leflore | Mississippi | 1 |  |
| 55 | Sunflower | Mississippi | 1 |  |
| 56 | Tunica | Mississippi | 1 |  |
| 57 | Washington | Mississippi | 1 |  |
| 58 | Westchester | New York | 1 |  |
| 59 | Allendale | South Carolina | 46 |  |
| 60 | Bamberg | South Carolina | 1 |  |
| 61 | Charleston | South Carolina | 1 |  |
| 62 | Lee | South Carolina | 51 |  |
| 63 | Williamsburg | South Carolina | 1 |  |
| 64 | Davidson | Tennessee | 1 |  |
| 65 | Haywood | Tennessee | 40 |  |
| 66 | Lake | Tennessee | 1 |  |
| 67 | Anderson | Texas | 1 |  |
| 68 | Bee | Texas | 1 |  |
| 69 | Frio | Texas | 1 |  |
| 70 | Madison | Texas | 65 |  |
| 71 | Walker | Texas | 1 |  |
| 72 | Willacy | Texas | 1 |  |
| 73 | Buckingham | Virginia | 75 |  |
| 74 | Greensville | Virginia | 1 |  |
| 75 | Sussex | Virginia | 50 |  |

Note: The rank is based on the sum of the posterior probabilities of a county being a COVID-19 infection hotspot and being a HIV infection hotspot

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Distribution of study variables** | **COVID-19 infection** | **HIV-infection** |
|  |  |  |  |
|  | **Mean (SD)** | **Relative Risk and** **(95% Credible Interval)** | **Relative Risk and** **(95% Credible Interval)** |
|  |  |  |  |
| % Black | **9.14% (14.59%)** | **1.133 (1.098—1.170)** | **1.543 (1.480—1.610)** |
| % Hispanic | **9.3% (13.86%)** | **1.305 (1.262—1.349)** | **1.175 (1.127—1.225)** |
| % aged 35 to 64 | **38.66% (2.98%)** | **0.960 (0.943—0.979)** | **1.102 (1.072—1.133)** |
| Socioeconomic deprivation a | **-1.85 (0.46)** | 1.022 (0.994**—**1.052) | **1.064 (1.024—1.106)** |
| Population density | **272.94 (1812.99)** | **0.972 (0.950—0.994)** | **1.056 (1.026—1.084)** |
| Chlamydia rate | **391.85 (277.08)** | **1.066 (1.043—1.087)** | **1.170 (1.132—1.209)** |
| % uninsured under 65 | **14.19% (6.14%)** | **1.112 (1.071—1.153)** | 1.031 (0.982**—**1.081) |
| Gini coefficient | **0.43 (0.04)** | **1.043 (1.025—1.062)** | **1.114 (1.084—1.145)** |
| % binge drinking | **16.95% (2.98%)** | **1.142 (1.106—1.177)** | **1.097 (1.052—1.144)** |

 aA principal components index of percent of population under age 25 with less than a high school diploma; median household income; percent of population 16 years and older unemployed; percent of people age 18 to 64 living in poverty.

***Appendix Table 2*: Multivariate results from Bayesian spatial Poisson regression analysis for factors associated with COVID-19 and HIV infection when implemented with a hierarchical model that accounts for potential clustering of counties within the 9 census regions. The Deviance Information Criterion (DIC) between this fitted model used for robustness check and the result without clustering is smaller than 2 (49650.35 VS 49648.99), suggesting that the models fitted the data equally well.**

**Notes. Details on Bayesian multivariate spatial statistical modeling**

Let ***Yik*** be the observed counts of COVID-19 (k=1) or HIV infection (k=2) at county i (i=1, 2, 3…, 3108). ***Yik*** is assumed to follow a Poisson distribution with expected mean ***mik\*Pik*** (Formula A1), where ***mik*** is the underlying infection rate and ***Pik*** is the total population at risk. If the HIV infection count is between 2 and 4 thus is suppressed, a Poisson distribution truncated to the range [2, 4] is used. Via a log link function, ***mik*** is further decomposed into (Formula A2): 1) ***αk*** that represents the underlying average infection rate over the whole country for outcome ***k***; 2) ***βm\*Xim***,where ***Xim*** is the *mth* (m=1,2,3,…,9) social, economic, behavioral, or clinical determinant at the *ith* county with a corresponding coefficient ***βm***, and 3) ***sik***, which is the spatial random effects term, representing spatial variations of the infection rates that are not captured by the covariates. In the case when census-region level variations are accounted for, Formula A2.1 with an additional random effects term, , is used instead, where census[i] is the census region the ith county belongs to. A non-informative flat prior was assigned to ***αk***(Formula A3). The prior for the coefficient ***βm*** was a vague Normal distribution with mean zero and variance 1000 (Formula A4). A multivariate Leroux prior was specified to the spatial random effects term **s** (Formula A5), where ***J*** is the total number of neighbors of county ***i***. Two counties sharing at least one vertex are considered as neighbors (a.k.a., first-order queen-based contiguity), a commonly used approach to define neighborhoods. ***wij***=1 if two counties ***i*** and ***j*** are neighbors, otherwise ***wij***=0. **Ω** represents a 2\*2 covariance matrix, which was assigned an inverse Wishart prior with a matrix parameter ***R*** (=**I2**, a 2\*2 identity matrix) and degree of freedom ***df*** (=2) (Formula A6) 1. A uniform prior restricted between 0 and 1 was specified to the spatial correlation parameter ***ρ***. An alternative, less informative prior was assigned to **Ω** with ***R***=0.01\***I2** 2. Almost identical results were obtained, suggesting that the results are not sensitive to hyperparameter prior specifications.

The models were fitted using WinBUGS v.1.4.3 3 via the R package *R2WinBUGS*. Two chains with diverging initial values were generated. Model convergence was assessed by visually checking the history and tract plots, and examining the Brooks-Gelman-Rubin (BGR) diagnostics. The model converges with a BGR value close to 1.0. The first 50,000 iterations, where the model converges, were discarded as burn-ins. We ran the model for another 50,000 iterations, resulting in a total of 100,000 samples for posterior inferences. The Deviance Information Criterion (DIC) was used to compare model-fitting. A model with a lower DIC or WAIC value fits the data better. Posterior predictive checks were used to evaluate model performance, that is, how well the fitted model predicts the observed data. A posterior predictive p-value close to 0.5 suggests that the replicated data is comparable to the observed data.

 (A1)

 (A2)

 (A2.1)

 (A3)

 (A4)

 (A5)

 (A6)

 (A7)

**References**

1 Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. 2014.

2 Martinez-Beneito MA, Botella-Rocamora P. Diesease mapping: From Foundations to Multidimensional Modeling. CRC Press, 2019.

3 Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat Comput* 2000; **10**: 325–37.