Distribution of Infant Mortality and Its Disparity by Socioeconomic Status

Jay H. Kim, Ph.D.
National Center for Health Statistics
Hyattsville, MD

1. Introduction

Infant death records have been linked to the corresponding birth certificates since 1995 by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention. Using the linked birth/infant death records from 2000 through 2006, infant mortality rates (IMRs) per 1,000 live births have been calculated (Table 1).

2. Objective

The focus of this study is to identify geographic differences by using cluster and hot/cold spot techniques of spatial statistics and plotting the differences on national maps by health service area (HSA)[1] using ArcGIS mapping program(ESRI, ArcGIS, 2008). IMR by sex and race are examined with regard to socioeconomic status (SES) of the mothers' residences.

HSAs are counties or groups of contiguous counties and are defined based on the travel pattern of hospital usage by Medicare patients (Makuc, 1991). The SES groups are constructed using 14 social and economic variables as reported by the 2000 Census using principal component method (Kim et al., 2009).

3. Data Source

Birth and infant mortality data for 2000-2006 used in this study have become available from the linked birth/infant death data files obtained by special request from the National Vital Statistics System of NCHS.

Socioeconomic indicators by county from the U.S. Census 2000 are merged to 805 HSAs.

4. Methods

*The findings in this paper are those of the author and do not necessarily represent the views of the NCHS/CDC.

[1] Geographic unit as defined by travel patterns between counties by Medicare beneficiaries for routine hospital care (Makuc et al., 1991).

4.1 Constructing Socioeconomic Status[2]

To develop a Socioeconomic Status Index (SES Index), a principal component method is applied to the 14 SES variables by using the SAS procedure Proc PRINCOMP (SAS Institute, Inc.). The first principal component which captured the largest amount (48%) of the variance of the linearly combined 14 SES variables is used to construct the SES subgroups. The procedure involves standardizing the original variables and then entering them into the formula of the first principal component for each HSA. It turns out that the resulting values of the SES indices range from 13.23 to -7.57. All the HSAs are then divided into five equal sized groups using their SES indices. The five groups are denoted as SES 1, SES 2, SES 3, SES 4 and SES 5, respectively, with SES 1 being the highest.

Three race groups are constructed based solely on mother's race without considering Hispanic origins – white, black and other. The other includes American Indian or Alaska Native (AIAN), Asian, and Native Hawaiian or Pacific Islander (API).

4.2 Calculating *Adjusted* Infant Mortality Rates

Infant mortality rates are calculated within the five SES groups for each year from 2000 to 2006 by sex and race. The rates are the number of deaths divided by the total number of live births expressed as per 1,000 live births. Subsequently the rates are adjusted to reflect the percentage of linked deaths among all births for each year. Table 2 shows the adjusted IMRs by SES and sex and race as calculated by the following formula:

$$\mathrm{IMR}_{ij} = (d_{ij}/b_{ij}) * 1{,}000$$

where d_{ij} = the number of infant deaths occurring in the i^{th} SES group in the j^{th} year, and b_{ij} = the total number of live births in i^{th} SES group, for the j^{th} year.

4.3 Identifying Clusters – Moran's I_i and Z-Score

Moran's I is a measure of spatial autocorrelation (Moran, 1950). According to his formula, the IMR of a "target" HSA is compared to the IMR of a neighboring HSA as each HSA becomes the target in turn using the ArcGIS software from ESRI that assigns weights to each HSA in a pair based on the distance between the two HSAs. For the identification of clusters of similar IMRs, local Moran's I_i compares each IMR value in a pair of HSAs to the mean value from all HSAs and is defined as

[2] Detailed descriptions on creating socioeconomic status (SES) groups are reported in Kim et al., 2007.

$$I_i = \frac{(X_i - \bar{X})}{s^2} \sum_j w_{ij}(X_j - \bar{X})$$

where X_j is the IMR for the neighboring HSA$_j$, X_i is the IMR for the target HSA$_i$, $\bar{X}$ is the mean of IMRs of all HSAs, W_{ij} is the weight of the target-neighbor with a value equal to the inverse of the distance between the target HSA$_i$ and its neighboring HSA$_j$, and s^2 is the variance of the IMRs from all HSAs.

A large positive value for Moran's I_i means that the HSA is surrounded by HSAs with similar rates, either high or low IMRs, while a negative value for I_i means the HSA is surrounded by HSAs with dissimilar rates.

The statistical significance of Moran's I_i is tested by a Z-score, $Z(I_i)$, which measures the neighborhood similarity not simply due to chance. Assuming the distribution of IMRs is random, the expected value of I_i, $E(I_i)$, is calculated as the negative of the sum of the weights for all HSA pairs divided by $n - 1$ where n is the total number of HSAs. That is,

$$E(I_i) = \frac{-\sum_j W_{ij}}{n-1}.$$

To measure the departure from the null hypothesis of a random distribution, that is, a distribution of cluster(s), a Z-score for Moran's I_i is computed as the difference between the expected value of I_i and the observed I_i divided by the standard deviation of I_i, that is,

$$Z(I_i) = \frac{I_i - E(I_i)}{\sqrt{V(I_i)}}$$

where

$$V[I_i] = E[I_i^2] - E[I_i]^2.$$

4.4 Identifying Hot and Cold Spots – Getis-Ord G_i and G_i Z-score

After clusters of HSAs with similar IMRs are identified, the clusters concentrated with very high or very low IMR values within a defined distance could be further classified. Clusters with very high Z-scores are defined as hot spots and those with very low Z-

scores are defined as cold spots. Those spots can be identified by using the Getis-Ord G_i statistic (Getis and Ord, 1996) as follows

$$G_i\,(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_j x_j}$$

where X_j is the IMR value for the j^{th} HSA within the target neighbors, and W_{ij} is the weight of X_j assigned to the target-neighbor with a value equal to the inverse of the distance (d) between two neighboring HSAs. To measure the departure from the null hypothesis of a random distribution, that is, a distribution of hot/cold spots, a Z-score for G_i is computed as the difference between the observed value of G_i and the expected value of G_i and divided by the standard deviation of G_i, that is,

$$Z(G_i) = \frac{G_i - E[G_i]}{\sqrt{V[G_i]}}$$

where $E[G_i]$ is the expected value of G_i, and is defined as[2]

$$E[G_i] = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}{n(n-1)}, \forall j \neq i$$

And

$$V[Gi] = E[Gi^2] - E[Gi]^2.$$

A high positive Z-score indicates a hot spot with statistically significant value of similar IMRs, and a low negative Z-score indicates a cold spot. A Z-score near zero indicates no apparent concentration of similar values.

5. Findings

Results indicate that the infant mortality rates remain relatively stable during the study period while continuing disparity among five socioeconomic and six race-sex groups. Male IMRs remain higher than female rates for all three races. Black males and females have consistently higher rates than white and other population (Table 2 and Figures 1-6).

The IMR maps of all race and sex of the U.S. display concentration of higher IMR areas where the SES is low, mostly in the southeast from southern Virginia through northern Florida to eastern Texas. The cold spots are located at the eastern Colorado through Nebraska, Iowa, Minnesota, the Dakotas, to Montana; some cold spots are also noticed in the northern California and southern Texas.

The patterns of geographic distribution of IMRs of white male and white female are similar – the hot spots are detected in the Appalachian Mountains spreading southwest to northern Louisiana with hot spots in Georgia and northern Florida for males.

The hot spots for black infants are detected in Wisconsin, Indiana and to Mississippi and the Appalachian mountains with a large region in southern Rocky mountains. Black female infants show hot spots in Illinois and Indiana region with northwestern Ohio and southern Rocky mountains.

References

Environmental Systems Research Institute. ArcGIS Software version 9.3.1, Redlands, CA, 2008.

Getis, A and Ord, JK. “Local Spatial Statistics: An Overview.” In Spatial Analysis: Modelling in a GIS Environment, edited by Longley P. and Batty M. Wiley, 1996.

Kim, JH, Wei, R and Gonzalez Jr, JF. Clusters and Hot Spots in the Distribution of Mortality: U.S.A., 2003-2005. Proceedings of Joint Statistical Meetings, 2009.

Makuc DM, Haglund B, Ingram DD, et al. Health Service Areas for the United States. National Center for Health Statistics. Vital Health Stat 2(112). 1991.

Moran, PAP. “Notes on Continuous Stochastic Phenomena,” Biometrika, 37, 17-33, 1950 and http://en.wikipedia.org/wiki/Moran%27s.

National Vital Statistics System. National Center for Health Statistics, Centers for Disease Control and Prevention. Linked Birth/Infant Death Data Sets. Various Years. Available at http://www.cdc.gov/nchs/linked.htm.

Appendix

Table 1. Number of Infant Deaths, Births and IMR*: 2000-2006, All Race and Sex.

Year	Infant Deaths	Births	IMR
2000	27,591	4,058,882	6.8
2001	27,268	4,026,036	6.8
2002	27,722	4,021,825	6.9
2003	27,710	4,090,007	6.8
2004	27,553	4,112,055	6.7
2005	28,021	4,138,573	6.8
2006	28,144	4,265,593	6.6

* Per 1,000 Live Births.

Table 2. IMR per 1,000 Live Birth by Race, Sex, Year and SES.

	2004			2005			2006		
Males	White	Black	Other	White	Black	Other	White	Black	Other
SES 1*	5.5	13.8	4.8	5.6	13.3	5.0	5.4	12.5	5.0
2	6.7	15.9	7.7	6.8	16.2	6.5	6.8	15.8	7.1
3	7.5	15.3	8.6	7.3	15.4	8.4	6.8	15.4	7.6
4	7.1	13.9	7.0	7.6	17.3	9.1	7.4	16.2	8.2
5	7.5	16.8	10.5	7.7	17.2	8.5	6.9	15.1	11.5

	2004			2005			2006		
Females	White	Black	Other	White	Black	Other	White	Black	Other
SES 1	4.5	11.2	4.4	4.5	11.1	4.5	4.6	10.9	3.9
2	5.4	12.7	4.5	5.3	12.5	6.1	5.3	13.2	5.5
3	6.1	12.5	7.0	5.9	11.6	6.8	5.4	13.5	6.0
4	5.9	12.6	6.6	6.3	15.0	5.7	5.4	12.0	7.4
5	6.0	12.5	7.4	6.0	12.4	6.9	5.7	13.1	5.8

*SES 1 being the highest.

Source: National Vital Statistics System, Linked Birth/Infant Death Data Sets. NCHS/CDC.

Figure 1. IMR by SES: Male, 2004.

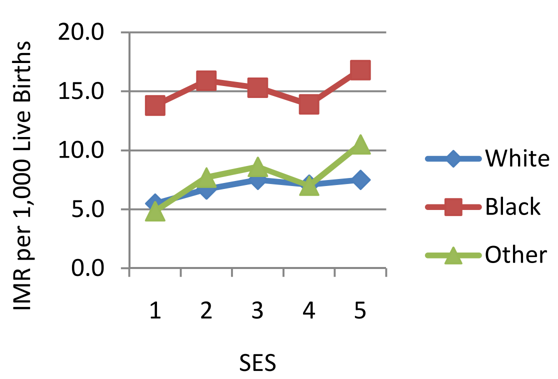


Figure 2. IMR by SES: Male, 2005.

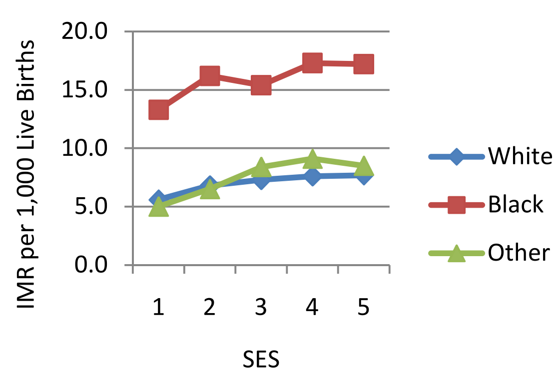


Figure 4. IMR by SES: Female, 2004.

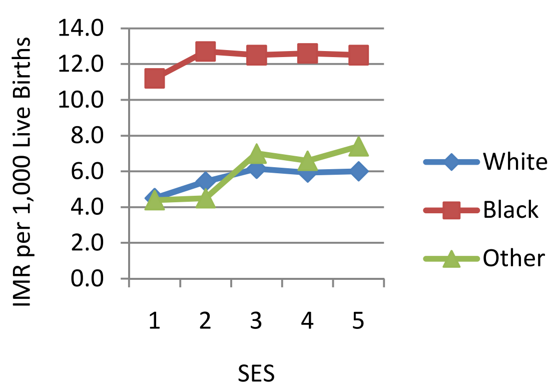


Figure 3. IMR by SES: Male, 2006.

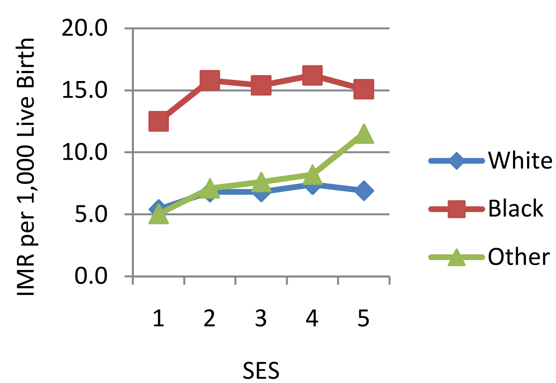


Figure 5. IMR by SES: Female, 2005.

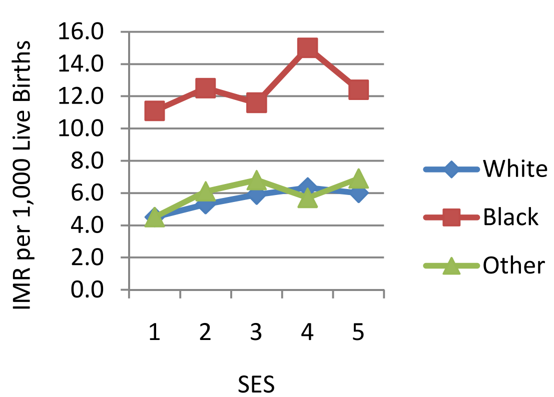
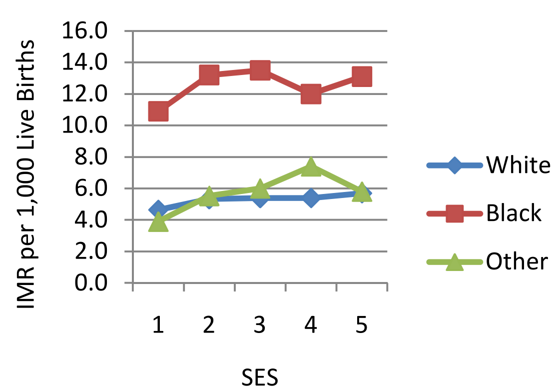
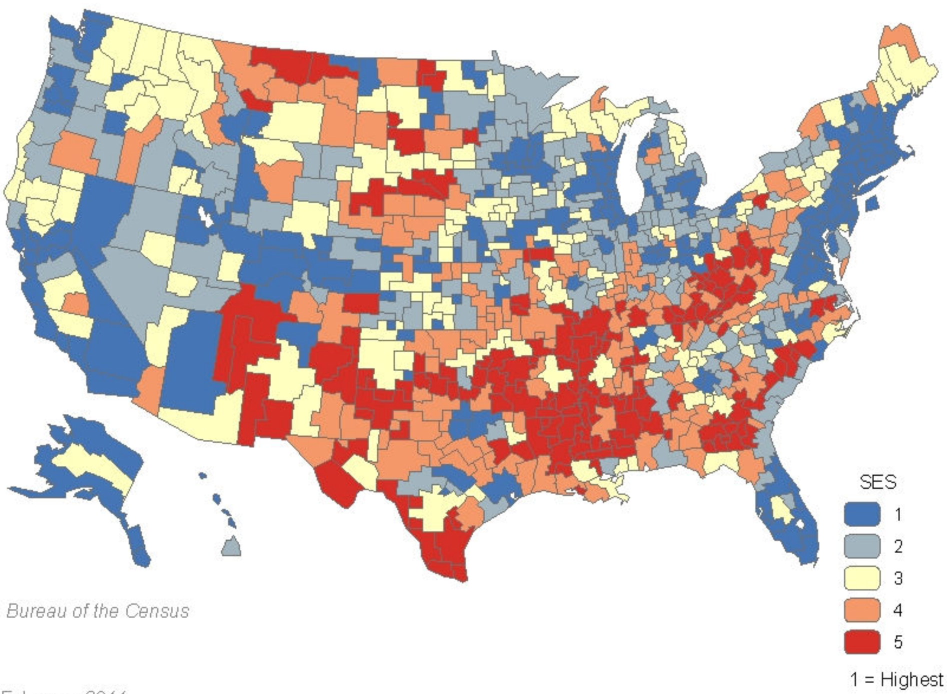


Figure 6. IMR by SES: Female, 2006.

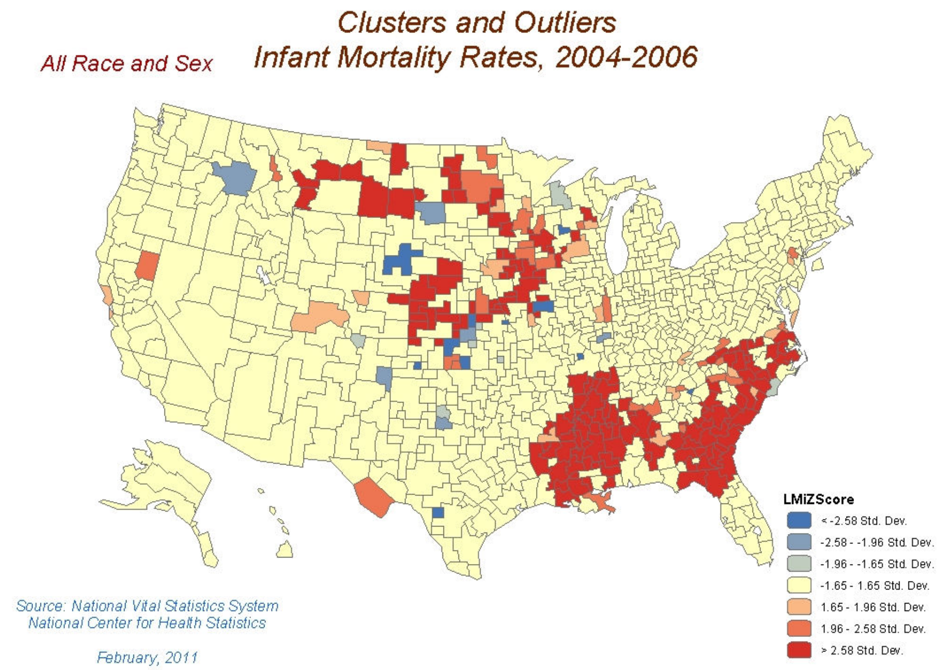


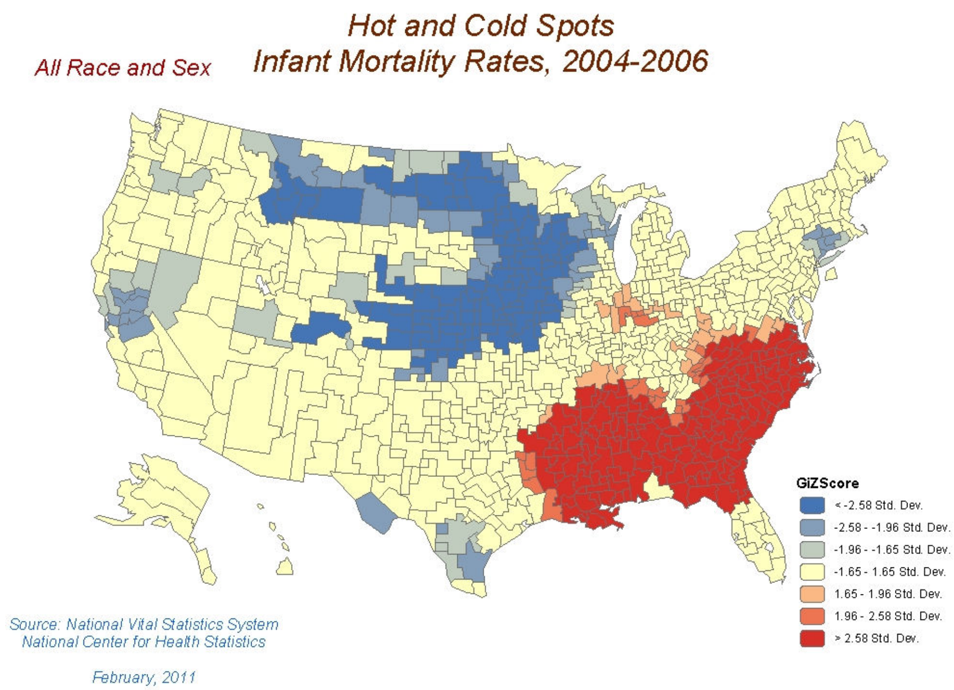
Socioeconomic Status by Health Service Area



Source: Bureau of the Census

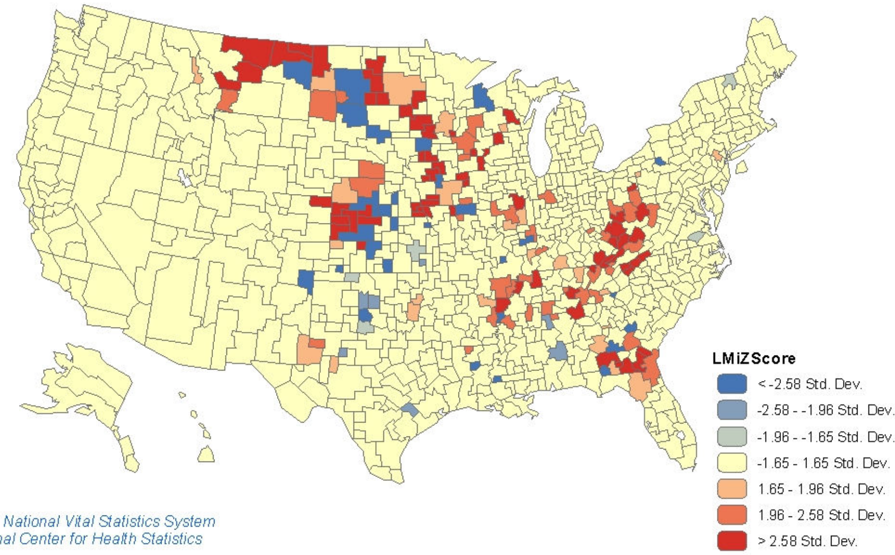
February, 2011


Clusters and Outliers
Infant Mortality Rates, 2004-2006
All Race and Sex
LMiZScore
< -2.58 Std. Dev.
-2.58 - -1.96 Std. Dev.
-1.96 - -1.65 Std. Dev.
-1.65 - 1.65 Std. Dev.
1.65 - 1.96 Std. Dev.
1.96 - 2.58 Std. Dev.
> 2.58 Std. Dev.
Source: National Vital Statistics System
National Center for Health Statistics
February, 2011


Hot and Cold Spots
Infant Mortality Rates, 2004-2006
All Race and Sex
GiZScore
< -2.58 Std. Dev.
-2.58 - -1.96 Std. Dev.
-1.96 - -1.65 Std. Dev.
-1.65 - 1.65 Std. Dev.
1.65 - 1.96 Std. Dev.
1.96 - 2.58 Std. Dev.
> 2.58 Std. Dev.
Source: National Vital Statistics System
National Center for Health Statistics
February, 2011

Clusters and Outliers

White Male Infant Mortality Rates, 2004-2006

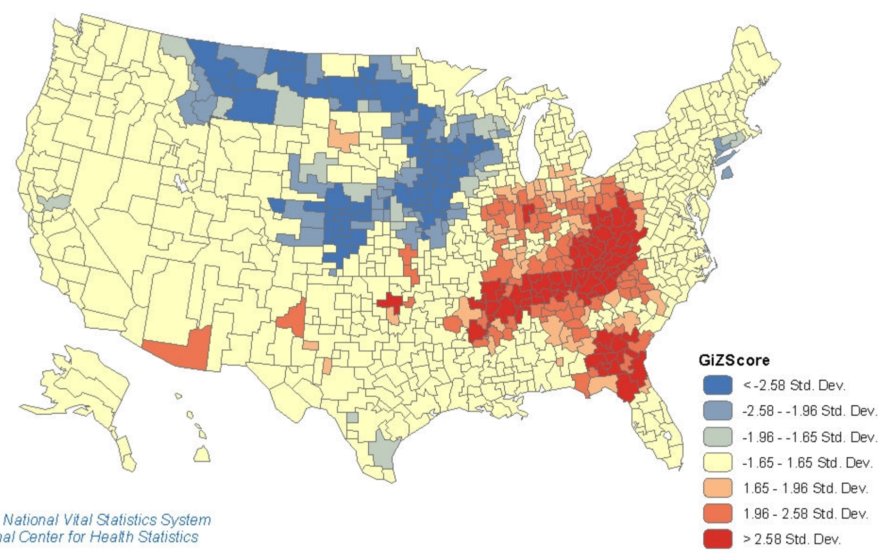


Source: National Vital Statistics System
National Center for Health Statistics

February, 2011

Hot and Cold Spots

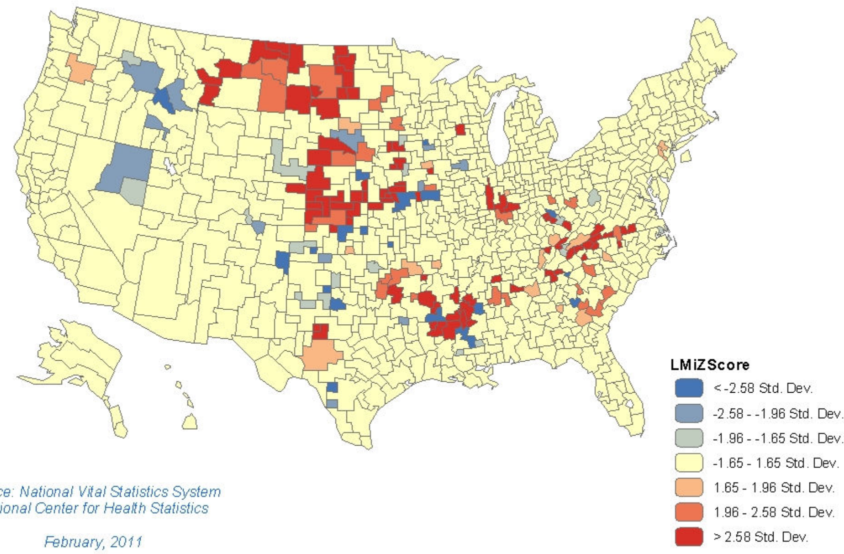
White Male Infant Mortality Rates, 2004-2006



Source: National Vital Statistics System
National Center for Health Statistics

February, 2011

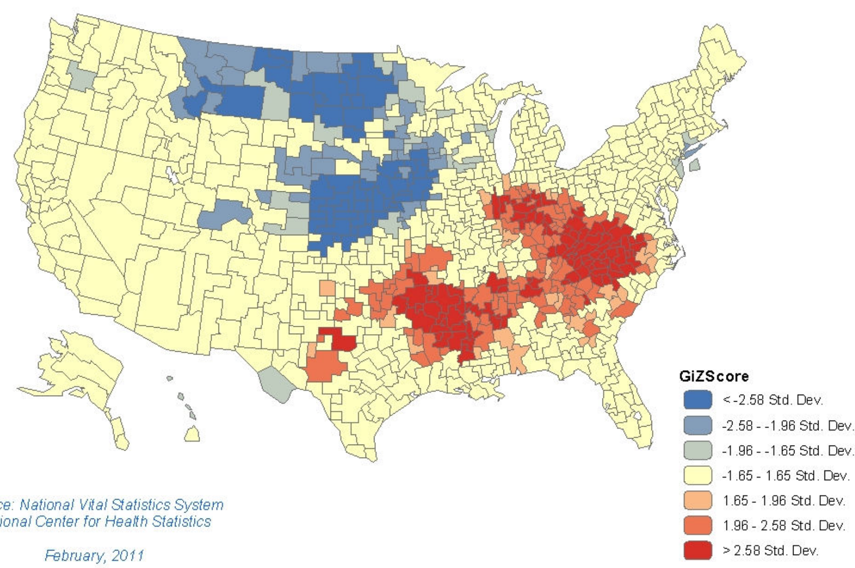
Clusters and Outliers
White Female Infant Mortality Rates, 2004-2006



Source: National Vital Statistics System
National Center for Health Statistics

February, 2011

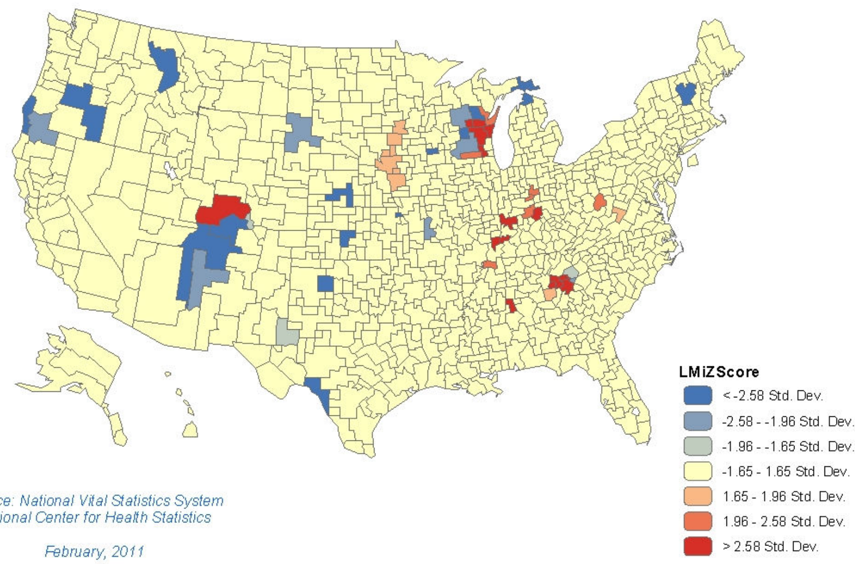
Hot and Cold Spots
White Female Infant Mortality Rates, 2004-2006



Source: National Vital Statistics System
National Center for Health Statistics

February, 2011

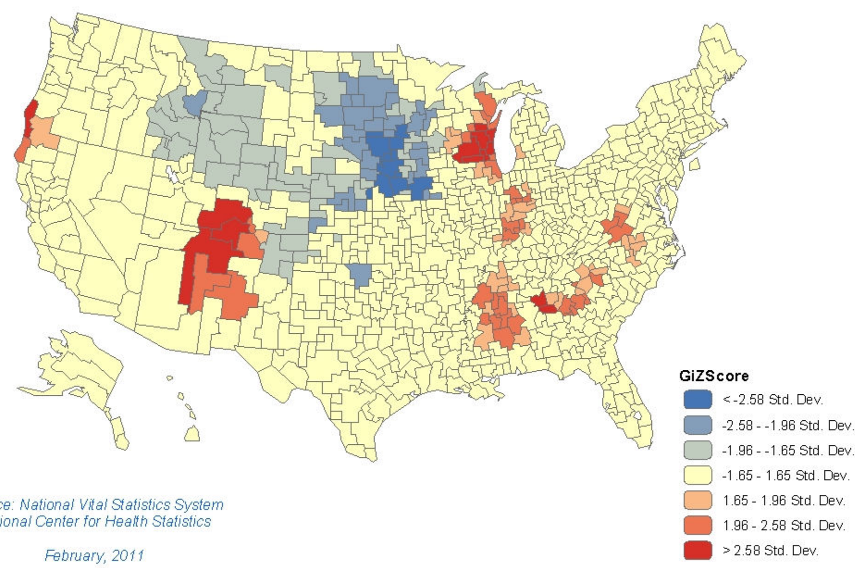
Clusters and Outliers
Black Male Infant Mortality Rates, 2004-2006



Source: National Vital Statistics System
National Center for Health Statistics

February, 2011

Hot and Cold Spots
Black Male Infant Mortality Rates, 2004-2006



Source: National Vital Statistics System
National Center for Health Statistics

February, 2011