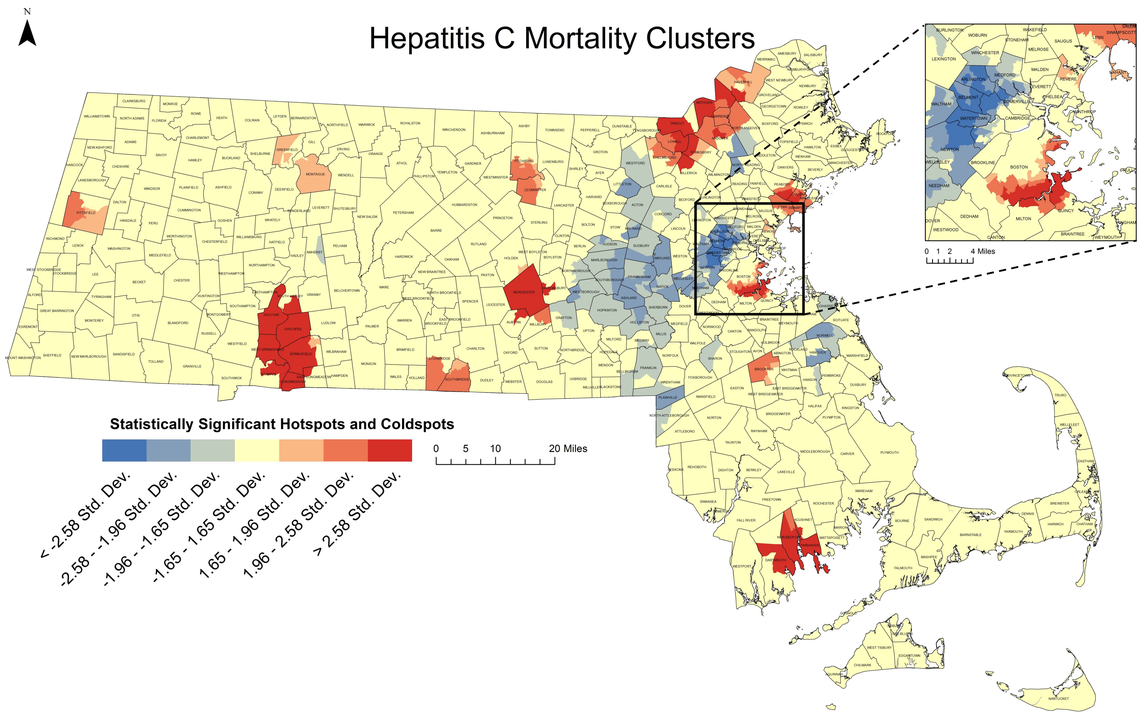
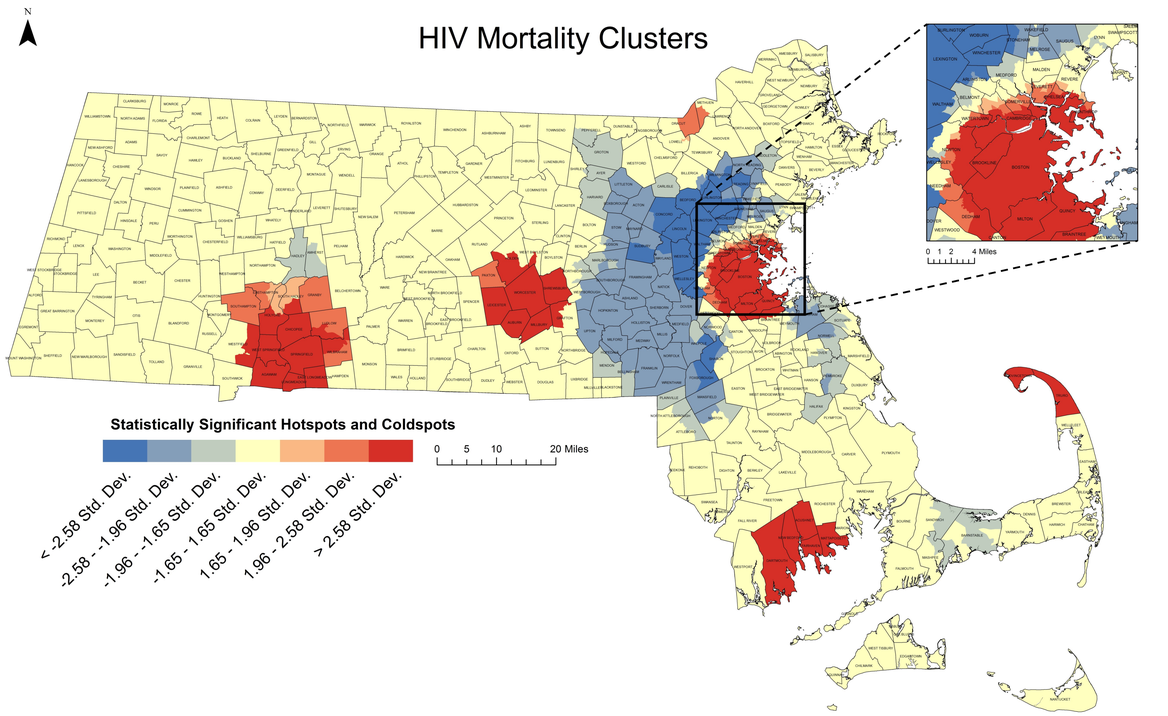
A Spatial Cluster Analysis of Massachusetts Infectious Disease Mortality, 2002-2011





Overview

Currently, trends in infectious disease mortality are not commonly examined in Massachusetts and there may be long term and cyclical patterns that are not being detected. The objective of this analysis is to examine spatial trends in infectious disease mortality in Massachusetts over a ten year period from 2002-2011 with a particular focus on HIV and Hepatitis C (HCV) in order to better inform public health resource allocation.

Background

HIV and Hepatitis C (HCV) are two sexually transmitted infections of great public health concern in Massachusetts. Deaths in HCV are increasing annually and while HIV in the United States in on a decline; it exists as a chronic condition for many people. Mortality rates of both of these conditions and their trends over time are not readily available.

Death certificates in Massachusetts allow for 1 underlying cause and 15 contributing causes of death. While most analysis conducted on mortality trends are conducted using only underlying cause, patterns may exist that are more visible using contributing cause as well.(Melamed & Sorvillo, 2009). This analysis makes use of both underlying and contributing cause of death to detect these patterns in HIV and HCV.

Hot spot cluster analysis is a useful method for determining where geographic clusters of disease exist(Getis & Ord, 1992). These analyses have been conducted to find clusters of disease and mortality occurrence in a variety of settings(Burra, Jerrett, Burnett, & Anderson, 2002; Gundogdu, 2010;). Stopka et. al. have developed a five step process to detect valid clusters of disease (Stopka, Krawczyk, Gradziel, & Geraghty, 2014). This method was used to determine geographic patterns of infectious disease mortality in this analysis.

Methods

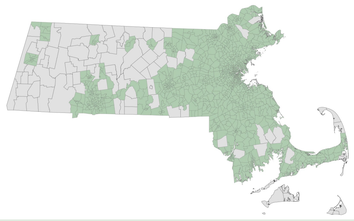
All deaths from 2002-2011 that featured HIV or HCV as an underlying or contributing cause of death was selected from the Massachusetts Death Certificate Database with census tract ID numbers. Census tract shapefiles were acquired from MassGIS. Mortality rates per 1,000 over the ten year period were calculated and aggregated to the tract level. These data were used to create descriptive choropleth maps to begin assessing for clusters in each disease condition.

5-Step Geoproccessing Approach

To examine potential hotspots or coldspots in disease mortality, Getis-Ord Gi* hotspot cluster analysis was used. However, before the analysis could be run, it was first necessary to select the parameters used in the analysis in an empirical manner. A 5-Step Geoproccessing approach developed by Stopka et al. was implemented to maximize the granularity and validity of the results.

Step 1: Analysis of Variation in Tract Sizes

In order to determine a proper spatial scale for running the hotspot analysis, all census tracts over 1.5 standard deviations of the Massachusetts census tract mean area were removed as they may distort the ideal sphere of influence in determining a cluster. Any tracts that did not share a border with at least 2 other tracts were also removed (pictured below).

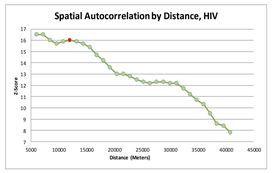


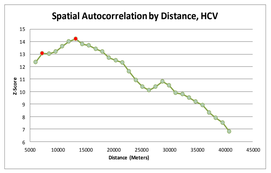
Step 2: Determining of Spatial Scale, Part A

In order to determine what distance should be used as a benchmark for cluster identification, the average and maximum distances between each tract and its 2 closest neighbors were calculated. The average distance was 2244 meters and the maximum distance was 9361 meters between each tract.

Step 3: Determination of Spatial Scale, Part B

For the greatest public health utility, it is necessary to calculate the smallest distance at which clustering of infectious disease mortality is intense. The Moran I spatial autocorrelation test is a statistical measure of the degree of the clustering of a given condition. To find the most adequate distance at which clustering is significant, two thirds of the maximum distance calculated in step 3 was used as a baseline distance for the Moran I test. 30 tests were run at increasing increments of half of the average distance calculated in the last step to find the smallest distance at which clustering (the z-score) peaks. For HIV this was at 12,000 meters. For HCV this was at 7,200 meters (pictured below).





Step 4: Accounting for the Larger Polygons

All of the census tracts were reintroduced to the map to be used in the creation of a spatial weights matrix, which takes into account a selected distance and a minimum number of neighbors in order to weight the relationship between each feature in the map. This matrix is used to adjust the final hotspot analysis.

Step 5: Hot Spot Analysis

In the final step a Getis-Ord Gi* Hot Spot Analysis was conducted for each disease condition weighed by each condition's spatial weights matrix. The Getis-Ord Gi* test calculates a Z-score for each feature indicating whether that feature exhibited clustering compared to the global mean. If there are enough high or low values in close vicinity to one another, then a hotspot or coldspot will be indicated.

Results

After performing hot spot cluster analysis following the 5-step Geoproccessing approach, there were 690 census tracts out of 1472 for HIV that were included in a statistically significant cluster ($p<.05$) and 440 census tracts that were included in a significant cluster for HCV. 500 of the HIV clustered census tracts were hotspots while 190 were coldspots.
330 of the HCV census tracts were in hotspots and 110 in coldspots.

Regions of significant clusters can be seen in the resultant maps above. Tracts in red displayed statistically significant hotspots based on their calculated z-score, while blue denotes statistically significant coldspots.

Discussion and Conclusions

The results of this analysis find several significant mortality hotspots and coldspots for Hepatitis C and HIV which can help inform resource allocation in the state. This approach is novel in that it makes use of both underlying and contributing causes of death to investigate clusters of mortality in an empirical way.

There exist several limitations of this study. Many practitioners receive little training in death certificate completion which can bias results. Additionally, when using multiple cause data, the decedent's death may not have had to do with the infectious disease in question. For example, if someone who was HIV positive died in a car accident, they would still be included in this analysis despite HIV not having contributed to their death.

As a result of this data using death data rather than incident data, the locations where people are currently dying of these conditions may differ from where current transmission is taking place. These clusters may be more indicative of people living with chronic HIV and HCV which is still of use to better target long-term treatment for these conditions.

Further research should take into account other socio-demographic characteristics such as race, income, gender, etc… as these factors may be correlated with high or low mortality rates for these conditions and may also exhibit geographic clustering. Hotspot cluster analysis when performed empirically presents a valuable resource for public health practitioners in identifying locations of high priority. Death certificate data used with a multiple cause approach also presents a valuable opportunity for characterizing trends in infectious disease.

References

Burra, T., Jerrett, M., Burnett, R. T., & Anderson, M. (2002). Conceptual and practical issues in the detection of local disease clusters: a study of mortality in Hamilton, Ontario. *Canadian Geographer / Le Géographe Canadien*, *46*(2), 160–171. doi:10.1111/j.1541-0064.2002.tb00737.x

Getis, A., & Ord, J. K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, *24*(3), 189–206. doi:10.1111/j.1538-4632.1992.tb00261.x

Gundogdu, I. B. (2010). Applying linear analysis methods to GIS-supported procedures for preventing traffic accidents: Case study of Konya. *Safety Science*, *48*(6), 763–769. doi:10.1016/j.ssci.2010.02.016

Melamed, A., & Sorvillo, F. J. (2009). The burden of sepsis-associated mortality in the United States from 1999 to 2005: an analysis of multiple-cause-of-death data. *Critical Care*, *13*(1), R28. doi:10.1186/cc7733

Stopka, T. J., Krawczyk, C., Gradziel, P., & Geraghty, E. M. (2014). Use of Spatial Epidemiology and Hot Spot Analysis to Target Women Eligible for Prenatal Women, Infants, and Children Services. *American journal of Public Health*, *104*(S1), S183–S189.

David Meyers | NUTR 231 | Spring 2014

Coordinate System: NAD 1983 Massachusetts State Plane

Data Sources: MassGIS; Division of Research and Epidemiology, Massachusetts Department of Public Health

Tufts University | School of Medicine
Public Health & Professional Degree Programs