Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Catherine M. Bulka Date

RESIDENCE PROXIMITY TO BENZENE RELEASE SITES IN GEORGIA AND NON-HODGKIN'S LYMPHOMA

By

Catherine M. Bulka

Master of Public Health

Global Epidemiology

_____

Dr. William McClellan, MD, MPH

Faculty Thesis Advisor

_____

Dr. Christopher Flowers, MD, MS

Field Advisor

RESIDENCE PROXIMITY TO BENZENE RELEASE SITES IN GEORGIA AND NON-HODGKIN'S LYMPHOMA

By

Catherine M. Bulka

Bachelor of Arts

Emory University

2010

Field Advisor: Dr. Christopher Flowers, MD, MS

Faculty Thesis Advisor: Dr. William McClellan, MD, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Epidemiology
2012

Abstract

RESIDENCE PROXIMITY TO BENZENE RELEASE SITES IN GEORGIA AND NON-HODGKIN'S LYMPHOMA

By Catherine M. Bulka

Background: The incidence rate of non-Hodgkin's lymphoma (NHL) has increased dramatically since the 1970s. This may be due to increased industrial production in the U.S., which has increased exposures to industrial chemicals. Benzene is one such chemical that is a known carcinogen and has been linked to various blood disorders. Increased risk of NHL has been observed in persons occupationally exposed to benzene, but the risk among persons living near benzene release sites has been less studied. The purpose of this study was to use data collected by the Georgia Comprehensive Cancer Registry (GCCR) to investigate the spatial patterns of NHL incidence and the association between NHL incidence and distance to benzene release sites.

Methods: We used population data from the U.S. Census and identified benzene release site locations from the EPAs Toxics Release Inventory. NHL cases were aggregated to census tracts. We performed Poisson regression on NHL incidence rates, using the mean distance between the tracts centroids and release sites as a marker of exposure. We mapped standardized incidence ratios by census tract to examine spatial patterns. Cluster analyses were conducted at the global, local, and focal levels.

Results: Local clusters of high standardized incidence ratios were identified in the metro-Atlanta area. Significant focal clustering of incidence rates was identified surrounding fourteen of the nineteen benzene release sites. The mean distance of census tract centroids from benzene release sites had a weak protective effect on NHL incidence, with an approximate 0.4% decrease in incidence for every mile the mean distance increased.

Conclusions: Increased mean distance of census tract centroid from release sites was associated with decreased NHL incidence. Clusters of NHL were significantly spatially associated with benzene release sites located in the metropolitan Atlanta area, but not with release sites in other areas of the state.

RESIDENCE PROXIMITY TO BENZENE RELEASE SITES IN GEORGIA AND NON-HODGKIN'S LYMPHOMA

By

Catherine M. Bulka

Bachelor of Arts

Emory University

2010

Faculty Thesis Advisor: Dr. William McClellan, MD, MPH

Field Advisor: Dr. Christopher Flowers, MD, MS

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Epidemiology
2012

ACKNOWLEDGEMENTS

Dr. William McClellan, I am sincerely grateful for your support and guidance on this project.

Dr. Christopher Flowers, your expertise and guidance are extremely valued.

Dr. Kevin Ward, thank you for your understanding and assistance.

Dr. Lance Waller, thank you for your guidance.

Thank you to my family - Mom, Dad, Nicole, Butterscotch, and Jack - for your unending love and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

B-NHL – B-cell non-Hodgkin's lymphoma

CLL/SLL – Chronic lymphocytic leukemia/small lymphocytic lymphoma

DLBCL – Diffuse large B-cell lymphoma

EPA – Environmental Protection Agency

FL – Follicular lymphoma

GCCR – Georgia Comprehensive Cancer Registry

GIS – Geographic Information Systems

IQR – Interquartile range

LISA – Local indicators of spatial autocorrelation

NHL – Non-Hodgkin's lymphoma

SEB - Spatial Empirical Bayes

SEER – Surveillance, Epidemiology, and End Results Program

SIR – Standardized incidence ratio

T-NHL – T-cell non-Hodgkin's lymphoma

TRI – Toxics Release Inventory

INTRODUCTION AND BACKGROUND

Non-Hodgkin's lymphomas (NHL) are a group of malignancies that arise in the lymphoid tissue (1). While the etiologies of most NHL cases are not well understood, the temporal trend regarding increased incidence is dramatic. In the 1970s and 1980s, incidence of NHL increased 3-4% annually. In the 1990s, incidence rates stabilized, but continued to increase by approximately 1-2% per year (1). While the AIDS epidemic may be partially responsible for this increasing incidence, it is possible that the rise in the number of cases has lagged slightly behind expanded industrial production in the U.S., suggesting occupational chemical exposures are risk factors for NHL (2). Of the occupational chemicals known to cause cancer, benzene has been consistently linked to hematological cancers (3). Benzene exposure has been shown to have toxic effects on the blood and bone marrow by reducing progenitor cell colony formation (4). Those with genetic variants in key metabolic enzymes may be particularly susceptible to this hematotoxicity. These toxic effects can occur at air levels of 1 ppm or less, suggesting that even low levels of benzene exposure can be harmful.

Non-Hodgkin's Lymphoma

Non-Hodgkin's lymphomas are neoplasms that affect lymphocytes, white blood cells that are part of the body's immune system (5). There are two main types of lymphocytes: B lymphocytes (B-cells) and T lymphocytes (T-cells). Normally, the B- and T-cells protect the body from infection. B-cells do this by making proteins called antibodies that attach to bacteria or viruses, acting as markers for destruction by other immune system cells. There are multiple types of T-cells, each with their own function, such as destroying any infected cells, releasing substances that attract other white blood cells that digest the infected cells, or boosting the activity of other immune system cells. Most of the cells that comprise lymphoid tissue are B- or T-cells. Lymphoid tissue is found in

various locations of the body, and thus lymphomas can develop in many different locations, often progressing through both solid tumor and circulating phases. Major sites of lymphoid tissue include the lymph nodes, spleen, thymus, adenoids, tonsils, digestive tract, and bone marrow.

The American Cancer Society (ACS) projected that of all new cancer cases in the U.S. during the year 2012, 4% could be attributed to NHL (6). This translates to an estimated 70,130 new cases of NHL in 2012. For the state of Georgia, ACS projects there will be 1,840 cases of NHL in 2012. Overall, NHL incidence rates are 40-70% higher among whites when compared to blacks. In terms of geographical distribution, NHL has historically been approximately 40% higher in urban areas than rural areas (1). For both males and females, the probability of developing NHL increases with age, but lymphoid cancers occur more frequently in males throughout all age groups (1). In many ways, NHL among adults differs greatly from NHL among children. Due to the short latency period of childhood NHL, it is likely that that the genetic aberrations that lead to NHL induction occur very early in life and possibly while in utero. Conversely, for adults, NHL is likely the result of lifestyle or environmental risk factors. Adulthood NHL is classified, staged, and treated differently than childhood NHL (10).

Benzene Toxicity

Upon entering the body, benzene is converted to metabolites in the liver and bone marrow (7). While most of these metabolites leave the body within 48 hours after exposure, both acute and chronic exposures can produce health effects. The severity of these effects is dependent on several factors, such as the amount of benzene to which the body is exposed, and the length of time exposed. Acute exposure to high levels of benzene can result in death, while acute exposure to low levels of benzene can result in dizziness, headache, tremors, confusion, rapid heart rate, and unconsciousness.

Typically, these effects do not last long after the body is no longer exposed. Long-term exposure to low levels of benzene through inhalation is harmful to the tissues that form blood cells, such as the bone marrow. This damage can disrupt the normal production of blood, as well as reduce the amount of important components in the blood, possibly causing anemia or excessive bleeding. When benzene exposure is excessive, the immune system's functionality can be impaired, reducing the body's ability to fight off infections. Long-term exposure to benzene through contaminated food and water is not well understood in humans, but animal studies have shown ingestion of benzene-contaminated food or water can damage the blood and the immune system, as well as cause cancer.

Numerous studies of mice have found associations between benzene exposure and the development of lymphomas (8). Biologically, benzene exposure is able to produce chromosomal aberrations and genetic changes necessary for NHL induction. Development of NHL can also occur as the result of chemotherapy and radiation treatments for a prior solid or hematologic cancer (9). It has been posited that benzene may act as an alkylating agent or as ionizing radiation, similar to these cancer treatments, in inducing NHL (8). In addition to inducing genetic changes, benzene has also been associated with epigenetic changes, including DNA methylation patterns commonly seen in cancerous cells, even at low doses (12).

Latency

The latency period for developing NHL as the result of an environmental exposure is largely unknown (11). For the development of secondary NHL following chemotherapy or radiation treatment of a prior solid or hematologic cancer, the median latency period is approximately 5 years, with a range of 2 to 15+ years. Among Hiroshima and Nagasaki atomic bomb survivors who were exposed to >100 rads of radiation, latency was

dependent on age at exposure. For those exposed to >100 rads before the age of 25, the average latency period was 9 years while older survivors had a longer average latency period of 14 years. For acute leukemia resulting from chronic low-dose exposure to benzene, the median latency period is between 15 and 20 years.

Epidemiologic Studies of Benzene Exposure and Non-Hodgkin's Lymphoma

An association between exposure to benzene and hematological effects was observed as early as the 1940s (13). In a 35-year cohort study of rubber workers from this time period, linear regression models demonstrated significant decreases in white and red blood cell counts, as well as hemoglobin, for workers exposed to benzene. More recent occupational studies have also shown similar associations. In a 1997 study of benzene-exposed workers in China, workers exposed to benzene for 10 or more years were 4.2 (95% CI = 1.1-15.9) times as likely to develop NHL than workers not exposed to benzene (14). Additionally, development of NHL was most strongly linked to exposures to benzene that had occurred at least 10 years before diagnosis when compared to recent exposures (between 1.5 to 10 years earlier), suggesting that benzene-related NHL be associated with longer induction periods.

Often, it is difficult to tease out the specific role benzene plays in the elevation of NHL risk, because workers exposed to benzene in an occupational setting tend to also be exposed to other potential carcinogens. In a systematic review of 43 case-control studies by Smith et al. that analyzed occupational benzene exposure and risk of NHL, only one study agreed that the only significant solvent exposure was benzene (20). This particular study by Rinsky et al. examined benzene exposure among rubber hydrochloride workers and leukemia, multiple myeloma, and NHL (21). While this study did observe significant elevations in the risks for leukemia and multiple myeloma, it failed to identify an association between benzene and NHL. In addition to the difficulties in identifying the

specific role of benzene among workers exposed to multiple potential carcinogens, the latency period of NHL, which is longer than the latency period for leukemia, may explain why the Rinsky study failed to observe an association with NHL but did with leukemia. Overall, the systematic review by Smith et al. supported a link between occupational benzene exposure and NHL with 93% of the 43 case-control studies identified showing an elevated risk of NHL, and 23 of these finding a significant association between probable exposure to benzene and NHL risk (20).

Given the occupational evidence supporting a link between benzene exposure and NHL, research is necessary to investigate associations between benzene exposure and NHL in non-occupational settings due to benzene releases into ambient air and surface waters. The purpose of this study was to determine if the incidence of non-Hodgkin's lymphoma at the census tract level between 1999-2008 in the state of Georgia varies with distance to the location of benzene release sites. To our knowledge, no previous studies have been conducted that specifically explore the relationship between distance to benzene toxic release sites and NHL incidence at the census tract level.

METHODS

We used three sets of secondary data in our analyses, including the location of benzene release sites in Georgia between 1988-1998 from the Environmental Protection Agency's Toxics Release Inventory, ten-year non-Hodgkin's lymphoma incidence from the Georgia Comprehensive Cancer Registry for 1999-2008, and population and demographic data from the United States Census from the year 2000.

Benzene Release Site Data

The Toxics Release Inventory provides information about releases of certain chemicals, including the quantities and media type of the releases (e.g., point source air emissions, surface water discharges, etc.) as well as their Cartesian coordinates, from a variety of sources such as manufacturing facilities, certain service business, and federal facilities (15). Launched in 1987, the TRI requires that facilities that meet certain thresholds must report their disposal or releases for listed toxic chemicals to EPA and to the state or tribal entity in which the facility is located annually. Between 1988 and 1998, 19 facilities in Georgia reported benzene releases to the TRI, with some of these facilities reporting benzene releases for multiple years during this period.

Census Tract Population and Demographic Data

The 22nd decennial census from 2000 asked a limited number of questions were asked of every person and housing unit in the United States, including sex, age, and race (16). Data regarding these variables were obtained for use in this analysis from Summary File 1. Additional questions were asked of a sample of persons and housing units, including data regarding the year moved into residence. These data were obtained from Summary File 3. Samples were generally comprised of 1-in-6 persons and housing units. All data were available down to the census tract level. At the time of the 2000 census, there were

1,618 census tracts within the state of Georgia. Population and demographic data were available for 1,616 of these tracts.

Census tracts are small, statistical subdivisions of counties with an average population of 4,000 persons throughout the country. Generally, census tracts have stable boundaries and were designed to have relatively homogenous demographic characteristics. However, between the 2000 and 2010 census, census tract boundaries in the state of Georgia changed so drastically that it was not possible to interpolate population and demographic characteristics during the period from 2001 to 2008 for this analysis. Thus this analysis was conducted under the assumption that these characteristics did not change during this time. Because accurate yearly population values were not available for each census tract, we divided NHL incidence data by ten calculating standardized incidence ratios and conducting regression analysis.

The U.S. Census compiles a value for median year moved into residence (MYMI) for a sample of individuals residing within census tracts. While this is an aggregate value, it was used to assess the general stability of the population. This variable provides a general idea of the length of time residents have been in their current home and thus, theoretically subject to the mean proximity to benzene release sites for this census tract. The average value for MYMI over all census tracts used in this investigation was 1989, which is during the period for potential benzene exposure (1988-1998) and allows for a theoretical period of ten years to develop NHL as our incidence data is for the years 1999-2008.

Non-Hodgkin's Lymphoma Data

The Georgia Comprehensive Cancer Registry (GCCR), part of the national SEER program, provided a case listing of all 12,716 incident NHL cases among adults ≥ 20 years residing in Georgia at the time of diagnosis, for the period 1999-2008. The cases

covered census tracts in 1,516 census tracts across the state. Cases were categorized by sex, race, and into fourteen age groups. Case addresses were manually checked for data entry errors that would have prevented accurate geocoding before they were geocoded to the census tract level. Cases that were geocoded with a match code of MA, MS, or M2 and a geolevel of B, G, or T were then aggregated to the census tract level. Match Codes and geolevels are variables of certainty provided for all records successfully geocoded using GIS. A match code of MA refers to a record matched to the Tele Atlas Address Points database; a match code of MS refers to a record matched to the Tele Atlas Street Address Ranges database; and lastly, a match code of M2 refers to a record matched to the ZIP+2 of the address (21). Geolevels refer specifically to census geography, with a to B referring to census blocks, a G referring to census block groups, and a T referring to census tracts. Cases without age, sex, or race information were excluded from all analyses. Additionally, because census data at the tract level in Georgia did not include population counts of persons whose race was categorized as "other" or "unknown," only cases categorized as "white," "black," "American Indian/Alaskan Native," and "Asian or Pacific Islander" were included. Incident cases were separated into five NHL subtypes (B-cell NHL, T-cell NHL, diffuse large B-cell lymphoma, chronic lymphocytic leukemia/small lymphocytic lymphoma, and follicular lymphoma) using ICD-O-3 codes based on the proposed WHO-based nested classification of malignant lymphoid neoplasms for epidemiologic research from InterLymph (17).

In order to standardize the NHL incidence rates from Georgia to national NHL incidence rates, SEER*Stat Version 7.05 (National Cancer Institute, Bethesda, MD) was used to access the SEER 13 Registries Database, which includes Atlanta, rural Georgia, Alaska, Connecticut, Detroit, Hawaii, Iowa, Los Angeles, New Mexico, San Francisco-Oakland, San Jose-Monterey, Seattle-Puget Sound, and Utah (18). Age-sex-race-specific

crude incidence rates were obtained for the period from 1999-2008 for NHL and NHL subtypes.

Geographic Data

We used Geographic Information System (GIS) through ArcGIS 10 to examine the spatial distribution of benzene release sites and standardized incidence ratios by census tract (ESRI, Redlands, CA). Census tract shapefiles were obtained from the U.S. Census Bureau's 2000 TIGER/Line files (20). Due to the polygon format of the census tract level data, centroids were calculated. GIS was then used to calculate the distances between these 1,616 census tract centroids to each of the 19 benzene release sites. To accurately calculate the distances between these points, maps were projected into the Georgia Statewide Lambert Conformal Conic projection.

Spatial Analyses

Descriptive spatial analysis was performed using ArcGIS. Choropleth maps were created to depict the standardized incidence ratios by census tract. Locations of benzene release sites were overlayed upon the SIR maps. Maps were visually assessed to determine the locations of high SIRs and identify any areas of potential clustering of NHL. The natural jenks method was used to classify SIR values.

Spatial Empirical Bayes (SEB) smoothing was performed on the SIR values using GeoDa 1.01 (Luc Anselin, Tempe, AZ). Due to the small geographical resolution of census tracts, sparse data can pose statistical problems. Spatial smoothing allows for census tracts to borrow strength from spatial neighbors to produce more stable estimates of SIRs. We defined neighbors as census tracts with first-order queen contiguity (census tracts with common borders or vertices). Choropleth maps of smoothed SIRs for NHL and NHL subtypes were created. SIR values were classified manually to reflect the same

structure as in the raw, or non-smoothed, SIR choropleth maps making the raw SIR maps and SEB-smoothed SIR maps visually comparable.

Global, local, and focal spatial analyses were used to examine the spatial structure of standardized incidence ratio and indirectly standardized incidence rate patterns and to identify significant clustering of elevated SIRs in the state. A global measure of spatial autocorrelation, the global Moran's I, was calculated for SIR (both raw and SEB-smoothed) patterns of NHL and NHL subtypes using GeoDa. This global Moran's I was used to identify the presence of any statistically significant spatial autocorrelation or clustering. To measure spatial autocorrelation at the local scale, a local Moran's I, also called LISA for Local Indicators of Spatial Autocorrelation, was calculated for SIR (both raw and SEB-smoothed) patterns of NHL and NHL subtypes using GeoDa. LISA cluster maps were also created to identify the specific locations of any statistically significant clusters of raw and SEB-smoothed SIRs using GeoDa. Clusters were categorized as "not significant," "high-high," "high-low," "low-high," and "low-low." Both the global and local spatial statistics based significance on 999 Monte Carlo simulations, and census tracts were weighted using queen contiguity.

The Lawson-Waller Score test was used to individually assess each of the 19 benzene release sites for focal clustering of NHL. Instead of using standardized incidence ratios to test for focal clustering, we used indirectly standardized incidence rates, calculated by multiplying the crude incidence rate by the standardized incidence ratio, to test the hypothesis that there is no spatial clustering near the benzene release sites (the focus). The alternative hypothesis is that the risk of NHL increases linearly with proximity to the release sites. Each census tract was scored for the difference between the observed and expected counts of NHL, weighted by inverse distance from each census tract centroid to the focus. By using indirectly standardized incidence rates, the effects of age, sex, and

race were eliminated. Scores were calculated and the standard normal distribution was used to estimate upper-tail p-values in ClusterSeer 2.3 (BioMedware, Ann Arbor, MI).

Statistical Analyses

In order to perform statistical analyses on the spatial data, the incidence data was aggregated to the census tract level then joined to the U.S. Census dataset. Data were then analyzed using SAS Version 9.3 (SAS Institute, Inc., Cary, NC). We conducted univariate analyses on indirectly standardized incidence ratios, mean census tract centroid distance to benzene release site, and demographic measures for each census tract. Pearson correlation analyses were performed to assess both the strength and direction of the relationships between the following variables: mean distance from census tract centroid distance to benzene release site, median year moved into residence, and the indirectly standardized incidence rates (both raw and SEB-smoothed) for NHL.

Poisson Regression Analyses

The mean distance from benzene release site in miles was calculated for the 1,616 census tracts. Indirectly standardized incidence rates of NHL were calculated by multiplying the SIR (both raw and SEB-smoothed) by the crude incidence rate of each census tract. Thus these rates were adjusted for age, sex, and race. The association between mean distance and the indirectly standardized incidence rate was determined using Poisson regression. Median year moved into residence was also assessed as a potential confounder and/or effect modifier. Incidence rate ratios along with 95% confidence intervals were obtained to determine the significance of association.

IRB Approval

The Emory University Institutional Review Board, the Winship Cancer Institute Institutional Review Board, and the Georgia Department of Public Health Institutional Review Board approved this study.

RESULTS

Of the 12,716 incident cases of NHL in our original dataset, 11,355 cases were able to be geocoded with a match code of MA, MS, or M2 and a geolevel of B, G, or T. From these, 11,323 had age, sex, and race information available and fit into the U.S. census race categories of "white," "black," "American Indian/Alaskan Native," or "Asian or Pacific Islander." Thus 89% of our original dataset that included all incident cases of NHL residing in Georgia at the time of diagnosis between the years of 1999 and 2008 was analyzed.

Spatial Distributions

We mapped standardized incidence ratios (indirectly standardized by age, sex, and race to national NHL and NHL subtype incidence rates) for NHL, NHL-B, NHL-T, DLBCL, FL, and CLL/SLL with point data for the benzene release sites. Two census tracts were not included in the analysis due to a lack of demographic data from the U.S. Census Bureau. Spatial Empirical Bayes smoothed SIRs were also mapped for NHL and all subtypes. Fourteen of the nineteen benzene release sites are in the greater metropolitan Atlanta area. The smoothed SIR maps show an apparent clustering of high SIRs for NHL and NHL-B in the metro-Atlanta area, but at a finer resolution many of the census tracts containing or surrounding these sites are actually less than or approximately equal to national incidence rates.

Spatial Analyses

The global Moran's I values were significant for NHL, NHL-B, DLBCL, and NHL-T raw SIRs. Moran's I values for all smoothed SIRs were significant. Standardized incidence ratios (both raw and SEB-smoothed) of non-Hodgkin's lymphoma were the most strongly spatially autocorrelated. LISA cluster maps (Figures 25 to 35) show the

locations of "hot-spots" (high-high clusters) and "cold-spots" (low-low clusters). Clustering of high SIRs appears to be located in the metro-Atlanta area for NHL and all subtypes, while clustering of low SIRs appears to be mostly in the southern region of the state.

Lawson-Waller score tests were statistically significant at the α=0.05 level for 15 of the 19 benzene release sites when using indirectly-standardized incidence rates and for 14 of the 19 benzene release sites when using spatial Empirical Bayes smoothed indirectly-standardized incidence rates. Interestingly, the sites with statistically significant positive scores for focal clustering are all located within the metropolitan Atlanta area, while all non-significant scores were located outside of metropolitan Atlanta.

Univariate Analyses

The results of univariate analyses for individual-level (case) data are presented in Table 1. The results of univariate analyses for census tract level variables are presented in Tables 2 and 3 and Figure 1. These results indicate the frequency of census tracts with no cases of NHL by the various subtypes. The more specific the NHL classification, the more census tracts have rates of zero. Means of census tract standardized incidence ratios for NHL are lower than one for NHL and all subtypes except DLBCL, indicating that Georgia has lower incidence of NHL, NHL-B, NHL-T, FL, and CLL/SLL than the national rates.

Correlations

We used Pearson correlation coefficients to assess the strength and relationship between indirectly standardized incidence rates and the mean distance from census tract centroid to benzene release site. Results are displayed in Tables 4 to 8. Pearson

correlation analysis was also performed to assess the median year moved into residence, included in this analysis as a potential confounder. Both the raw and smoothed incidence rates were weakly negatively correlated with mean distance to benzene release site for NHL and NHL subtypes except CLL/SLL, and these results were statistically significant at the $\alpha = 0.05$ level. Median year moved into residence was not significantly correlated with any variable.

Poisson Regression Analyses

No statistically Poisson coefficients emerged for the interaction term between mean distance to benzene release site and MYMI, so this variable was dropped. Additionally, comparing gold-standard models to models that did not include MYMI suggested that it was not a confounder in the association between the mean distance and the raw indirectly standardized incidence rate or in the association between the mean distance and the SEB-smoothed indirectly standardized incidence rate for NHL and NHL subtypes.

For NHL and all subtypes, with the exception of CLL/SLL, mean distance from benzene release site had a significant, but weak, protective effect on raw indirectly standardized incidence. Results were similar for SEB-smoothed indirectly standardized incidence. The regression models for the SEB-smoothed data fit better than the models for the raw indirectly standardized data. While the protective effect of mean distance from benzene release site was weak, with an approximate 0.4% decrease in the NHL incidence rate for every mile the mean distance increases, this translates to an approximate decrease of 32.4% for census tracts with a mean distance of 99.5 miles (the average value at the census tract level) from benzene release sites. Results are displayed in Tables 9 and 10.

DISCUSSION

We examined residential proximity to benzene release sites and risk of NHL in Georgia using census tract level data. We observed a weak protective effect of a 0.4% decrease in NHL incidence for each mile the mean distance between benzene release sites and the census tract centroid increased. In addition, cluster analyses identified several hot spots and cold spots for NHL and NHL subtypes throughout the state. The metropolitan Atlanta area was almost always identified as a hot spot, while other urban areas, namely Augusta and Savannah, were also sometimes implicated. Cold spots were most often located in the southern half of the state. A test of focal clustering identified significantly increased NHL incidence surrounding 15 of the 19 benzene release sites analyzed when using rates that had not been smoothed and 14 of the 19 benzene release sites when using smoothed rates. Only one benzene release site identified as having increased NHL incidence surrounding it was located outside of the metropolitan Atlanta area; this site, number 16, was instead located near Savannah, GA. However, this site was not significant when using smoothed rates. These results suggest that focal clustering is present around benzene release sites located in the metropolitan Atlanta, but this phenomenon is not observed in other areas of the state.

While the results supported our hypothesis, there are two alternative hypotheses that may also explain these results. Firstly, because NHL incidence was higher than expected in urban areas of the state after accounting for age, sex, and race, it is possible that these areas have higher access to health care and better cancer surveillance than rural areas, causing more NHL cases to be detected. However, GCCR has been collecting data on all cancer cases diagnosed amongst Georgia residents since 1995. A 2002 evaluation of GCCR determined that the system detects 97.6% of all cancer cases in the state, with only 2.4% detected by death certificate only (22). Thus GCCR accurately portrays the occurrence of NHL in the state and thus this alternative hypothesis does not appear to

explain the results of this analysis. The second alternative hypothesis that may explain the results observed is that populations living in urban areas are exposed to some unmeasured factor or an unaccounted for confounder. Besides point source industrial releases of benzene, there are numerous sources of exposure. Major sources of exposure include mainstream smoke from cigarettes and auto exhaust (23). It is possible that these exposures are higher in urban areas and thus explain the observed results.

STRENGTHS AND WEAKNESSES

Strengths of this study include the use of publicly available data from EPA and the U.S. Census Bureau to conduct a descriptive study. Additionally, this is the first study of its kind to use a combination of spatial cluster statistics, statistical modeling, and visual representation of data through GIS to analyze distance from benzene release sites and incidence of non-Hodgkin's lymphoma in Georgia. Spatial smoothing of NHL incidence data addressed differences in census tract population size that resulted in variance instability and spurious outliers. Lastly, by standardizing the data indirectly, we eliminated the effects of age, sex, and race on NHL incidence, in addition to allowing for comparisons of NHL incidence to be made between Georgia and other SEER 13 registries.

Limitations of this study include the use of data aggregated to the census tract level. Furthermore, it is important to note when using TRI data that the presence of chemical releases into the environment is not sufficient to determine personal exposure, or to calculate potential risks to human health, especially because many facilities limit contamination and potential exposure to humans by managing the chemicals in certain ways. Distance was used as a proxy for exposure, but was calculated using census tract centroids, thus findings may not hold true at the individual level. Another weakness of this study was the lack of temporal analyses; we did not study temporal or space-time clustering. However, we attempted to control for time and latency by only including benzene release sites from 1988 to 1998 and using incidence and population data from 1999 to 2008, as well as assessing the median year moved into residence as a potential confounder at the census tract level. Additionally, the amount of benzene released at the identified sites was not assessed, so any dose-response relationship was not studied.

The use of U.S. Census data as denominators for incidence rates may have also introduced bias to this study. Race was recoded into four major categories, but in the

2000 census, respondents could have chosen more than one racial category. This may have affected population estimates and thus could have affected calculations for NHL incidence standardized by race, age, and sex. Finally, because of changes in census tract boundaries between 2000 and 2010, we were unable to interpolate population estimates for annual incidence rates and thus based all calculations on population estimates from only the year 2000. Any significant changes in the population of Georgia between 1999 and 2008 may have affected NHL incidence rates.

We used Poisson regression to statistically model NHL incidence as a function of mean distance from benzene release sites. An assumption of Poisson regression is the independence of observations. However, using mean distance from benzene release sites as an explanatory variable by nature violates this assumption. In addition, any unmeasured confounders are likely also spatially correlated. Residuals may also be spatially autocorrellated. We did not address this in our analyses.

FUTURE DIRECTIONS

This study identified mean distance from benzene release sites as a risk factor for NHL at the census tract level. A future direction for this study should examine this association using individual-level geographic data. Temporal trends and dose-response relationships between benzene exposure and NHL also warrant further investigation. Instead of Poisson regression, spatial regression analyses should be conducted on this data in order to account for spatial dependence between observations. Future directions for this research should also evaluate other potential sources of benzene exposure in Georgia that may explain the higher incidence of NHL observed in urban areas of the state. It is crucial to further investigate the etiology of non-Hodgkin's lymphoma as the incidence continues to increase.

REFERENCES

(1) Müller, Antonia M. S., et al. "Epidemiology of non-Hodgkin's Lymphoma (NHL): trends, geographic distribution, and etiology." *Annals of Hematology*. 9 October 2004. Volume 84: 1-12.

(2) De Roos, A., et al. "Residential Proximity to Industrial Facilities and Risk of Non-Hodgkin Lymphoma." *Environmental Research*. 1 January 2010. Volume 110 (1): 70-78.

(3) Smith, Martyn T. "Advances in Understanding Benzene Health Effects and Susceptibility." *Annual Review of Public Health*. 4 December 2010. Volume 31: 133-148.

(4)Qing, Lan, et al. "Hematotoxicity in Workers Exposed to Low Levels of Benzene. *Science*. 3 December 2004. Volume 306 (5702): 1774.

(5) American Cancer Society. "Non-Hodgkin Lymphoma." *Detailed Guide*. 26 January 2012. Web. 20 March 2012. http://www.cancer.org/acs/groups/cid/documents/webcontent/003126-pdf.pdf

(6) American Cancer Society. *Cancer Facts & Figures 2012*. Atlanta, GA: American Cancer Society. 2012.

(7) Agency for Toxic Substances and Disease Registry (ATSDR*). "*Toxicological Profile for Benzene." Atlanta, GA: U.S. Department of Health & Human Services, US Public Health Service. 1997.

(8) Smith, Martyn T., et al. "Benzene Exposure and Risk of Non-Hodgkin Lymphoma." *Cancer Epidemiology, Biomarkers, & Prevention*. 2 March 2007. Volume 16: 385-391.

(9) Krishnan, Biju and Morgan, Gareth. "Non-Hodgkin Lymphoma Secondary to Cancer Chemotherapy." *Cancer Epidemiology, Biomarkers, & Prevention*. 19 March 2007. Volume 16: 377-380.

(10) Murphy, Sharon B. "Childhood Non-Hodgkin's Lymphoma." *The New England Journal of Medicine*. 28 December 1978. Volume 299: 1446-1448.

(11) Weisenburger, Dennis D. "Pathological Classification of Non-Hodgkin's Lymphoma for Epidemiological Studies." *Cancer Research*. 1 October 1992. Volume 52 (supplement): 5456-5462.

(12) Bollati, Valentina, et al. "Changes in DNA Methylation Patterns in Subjects Exposed to Low-Dose Benzene." *Cancer Research*. 5 February 2007. Volume 67: 876-880.

(13) Kipen, Howard M., et al. "Hematologic Effects of Benzene: A Thirty-Five Year Longitudinal Study of Rubber Workers." *Toxicology and Industrial Health*. 1 January 1988. Volume 4:411-430.

(14) Hayes, Richard B., et al. "Benzene and the Dose-Related Incidence of Hematologic Neoplasms in China." *Journal of the National Cancer Institute*. 9 May 1997. Volume 89: 1065-1071.

(15) Environmental Protection Agency. *The Toxics Release Inventory (TRI) and Factors to Consider When Using TRI Data*. Web. 22 March 2012. http://www.epa.gov/tri/triprogram/FactorsToConPDF.pdf

(16) U.S. Census Bureau. "Introduction to Census 2000 Data Products." June 2001. Web. 22 March 2012. http://www.census.gov/prod/2001pubs/mso-01icdp.pdf

(17) Morton, Lindsay M., et al. "Proposed classification of lymphoid neoplasms for research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph)." *Blood*, 27 March 2007. Volume 110: 695-708.

(18) Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 13 Registries Research Data, Nov 2010 Sub (1973-2008) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969-2009 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2011, based on the November 2010 submission.

(19) U.S. Census Bureau. "Census 2000 TIGER/Line® Files." *TIGER/Line Shapefiles and TIGER/Line Files*. 5 October 2001. Web. 9 January 2012. http://www.census.gov/geo/www/tiger/tiger2k/tgr2000.html

(20) Smith, Martyn T., Jones, Rachael M., and Smith, Allan H. "Benzene Exposure and Risk of Non-Hodgkin Lymphoma." *Cancer Epidemiology, Biomarkers & Prevention,* March 2007. Volume 16(3): 385-391.

(21) Rinsky, R., Hornung, R., Silver, S. and Tseng, C. "Benzene exposure and hematopoietic mortality: A long-term epidemiologic risk assessment." *American Journal of Industrial Medicine,* 18 November 2002. Volume 42(6): 474–480.

(22) Clarkson, L. "An Evaluation of the Georgia Comprehensive Cancer Registry: Improving an Established System." *Georgia Department of Human Resources, Division of Public Health*, February 2007.

(23) Wallace, L. "Major sources of benzene exposure." *Environmental Health Perspectives,* July 1989. Volume 82: 165-169.

Table 1. Demographics of study population

Variable	NHL (N=11,323) N (%)	NHL-B (N=8,925) N (%)	DLBCL (N=3,851) N (%)	FL (N=2,171) N (%)	NHL-T (N=1,489) N (%)	CLL/SLL (N=765) N (%)
Age Group						
20-24	119 (1.05)	85 (0.95)	51 (1.32)	8 (0.37)	19 (1.28)	1 (0.13)
25-29	179 (1.58)	123 (1.38)	80 (2.08)	14 (0.64)	39 (2.62)	1 (0.13)
30-34	284 (2.51)	200 (2.24)	115 (2.99)	38 (1.75)	57 (3.83)	3 (0.39)
35-39	432 (3.82)	304 (3.41)	181 (4.70)	62 (2.86)	83 (5.57)	9 (1.18)
40-44	630 (5.56)	465 (5.21)	227 (5.89)	117 (5.39)	106 (7.12)	19 (2.48)
45-49	750 (6.62)	573 (6.42)	240 (6.23)	160 (7.37)	127 (8.53)	37 (4.84)
50-54	1030 (9.10)	831 (9.31)	323 (8.39)	260 (11.98)	140 (9.40)	66 (8.63)
55-59	1161 (10.25)	947 (10.61)	343 (8.91)	278 (12.81)	150 (10.07)	110 (14.38)
60-64	1215 (10.73)	982 (11.00)	352 (9.14)	270 (12.44)	144 (9.67)	94 (12.29)
65-69	1253 (11.07)	1027 (11.51)	424 (11.01)	238 (10.96)	129 (8.66)	99 (12.94)
70-74	1268 (11.20)	1045 (11.71)	440 (11.43)	249 (11.47)	137 (9.20)	108 (14.12)
75-79	1275 (11.26)	1024 (11.47)	436 (11.32)	233 (10.73)	126 (8.46)	109 (14.25)
80-84	954 (8.43)	745 (8.35)	356 (9.24)	149 (6.86)	123 (8.26)	70 (9.15)
85+	773 (6.83)	574 (6.43)	283 (7.35)	95 (4.38)	109 (7.32)	39 (5.10)
Female	5274 (46.58)	4164 (46.66)	1779 (46.20)	1106 (50.94)	668 (44.86)	311 (40.65)
Male	6049 (53.42)	4761 (53.34)	2072 (53.80)	1065 (49.06)	821 (55.14)	454 (59.35)
White	8991 (79.40)	7272 (81.48)	3034 (73.78)	1919 (88.39)	1031 (69.24)	583 (76.21)
Black	2208 (19.50)	1551 (17.38)	757 (19.66)	233 (10.73)	444 (29.82)	179 (23.40)
American Indian/Alaskan Native	4 (0.04)	4 (0.04)	1 (0.03)	2 (0.09)	0 (0.00)	1 (0.13)
Asian or Pacific Islander	120 (1.06)	98 (1.10)	59 (1.53)	17 (0.78)	14 (0.94)	2 (0.26)

Table 2. Census tract level demographics

Variables	N=1,616 Mean (SD)
Age*	
20-24	372.5 (334.1)
25-29	372.9 (290.5)
30-34	388.0 (298.3)
35-39	417.8 (318.3)
40-44	394.4 (279.9)
45-49	347.2 (224.7)
50-54	308.6 (190.6)
55-59	229.3 (134.1)
60-64	174.7 (97.8)
65-69	145.0 (83.2)
70-74	122.1 (72.0)
75-79	96.7 (60.7)
80-84	64.0 (44.3)
85+	54.0 (50.9)
Female*	1799.2 (1005.6)
Male*	1655.4 (972.2)
White*	2427.0 (1894.7)
Black*	941.4 (1009.9)
American Indian/Alaskan Native*	9.5 (8.9)
Asian or Pacific Islander*	39.8 (81.5)
Mean MYMI*	1989.3
Mean distance from benzene release site (miles)	99.5 (55.2)

*From 2000 U.S. Census

Table 3. Census tract level variables

Variables	NHL N=1,616	NHL-B N=1,616	DLBCL N=1,616	FL N=1,616	NHL-T N=1,616	CLL/SLL N=1,616
Tracts with cases	1,516 (93.8%)	1,477 (91.4%)	1,274 (78.8%)	1,013 (62.7%)	866 (53.6%)	562 (34.8%)
Mean incidence (SD)*	21.17 (23.08)	20.45 (29.66)	93.84 (143.56)	7.38 (45.24)	5.96 (21.09)	1.02 (3.59)
(IQR 25%-75%)	(6.27-28.56)	(4.85-27.82)	(8.19-123.66)	(0.00-7.63)	(0.00-5.59)	(0.00-0.78)
Mean SEB-smoothed incidence (SD)*	17.80 (12.65)	16.22 (13.09)	62.53 (62.50)	3.72 (4.99)	2.69 (4.84)	0.36 (0.80)
(IQR 25%-75%)	(8.75-24.81)	(7.46-22.54)	(16.88-91.68)	(0.00-5.52)	(0.00-3.81)	(0.00-0.46)
Mean SIR (SD)	0.84 (0.52)	0.99 (0.84)	8.51 (7.34)	0.87 (1.21)	0.80 (1.08)	0.20 (0.51)
(IQR 25%-75%)	(0.52-1.09)	(0.53-1.34)	(3.07-12.57)	(0.00-1.36)	(0.00-1.28)	(0.00-0.33)
Mean SEB-smoothed SIR (SD)	0.84 (0.26)	0.99 (0.38)	8.61 (3.58)	0.87 (0.39)	0.78 (0.42)	0.20 (0.13)
(IQR 25%-75%)	(0.68-1.01)	(0.78-1.20)	(6.68-10.45)	(0.65-1.09)	(0.54-1.02)	(0.11-0.28)

*Incidence per 100,000, indirectly standardized by age, sex, and race for 1999-2008

Table 4. Bivariate correlations for NHL

		Pearson Correlation Coefficients, N = 1,616			
		Incidence	**Smoothed incidence**	**Mean distance**	**MYMI**
Incidence	*Correlation coefficient*	1.0000		-0.1546	-0.0032
	P-value			<0.001	0.8970
Smoothed incidence	*Correlation coefficient*		1.0000	-0.2766	-0.0027
	P-value			<0.001	0.9136
Mean distance	*Correlation coefficient*	-0.1546	-0.2766	1.0000	0.0002
	P-value	<0.001	<0.001		0.9948
MYMI	*Correlation coefficient*	-0.0032	-0.0027	0.0002	1.0000
	P-value	0.8970	0.9136	0.9948	

Table 5. Bivariate correlations for NHL-B

		Pearson Correlation Coefficients, N = 1,616			
		Incidence	**Smoothed incidence**	**Mean distance**	**MYMI**
Incidence	*Correlation coefficient*	1.0000		-0.1144	0.0259
	P-value			<0.001	0.2979
Smoothed incidence	*Correlation coefficient*		1.0000	-0.1797	0.0354
	P-value			<0.001	0.1548
Mean distance	*Correlation coefficient*	-0.1144	-0.1797	1.0000	0.0027
	P-value	<0.001	<0.001		0.9126
MYMI	*Correlation coefficient*	0.0259	0.0354	0.0027	1.0000
	P-value	0.2979	0.1548	0.9126	

Table 6. Bivariate correlations for DLBCL

		Pearson Correlation Coefficients, N = 1,616			
		Incidence	**Smoothed incidence**	**Mean distance**	**MYMI**
Incidence	*Correlation coefficient*	1.0000		-0.0944	0.0192
	P-value			0.0001	0.4403
Smoothed incidence	*Correlation coefficient*		1.0000	-0.1688	0.0169
	P-value			<0.0001	0.4963
Mean distance	*Correlation coefficient*	-0.0944	-0.1688	1.0000	0.0027
	P-value	0.0001	<0.0001		0.9126
MYMI	*Correlation coefficient*	0.0192	0.0169	0.0027	1.0000
	P-value	0.4403	0.4963	0.9126	

Table 7. Bivariate correlations for FL

		Pearson Correlation Coefficients, N = 1,616			
		Incidence	**Smoothed incidence**	**Mean distance**	**MYMI**
Incidence	*Correlation coefficient*	1.0000		-0.0302	0.0066
	P-value			0.2251	0.7908
Smoothed incidence	*Correlation coefficient*		1.0000	-0.0719	0.0264
	P-value			0.0038	0.2884
Mean distance	*Correlation coefficient*	-0.0302	-0.0719	1.0000	0.0027
	P-value	0.2251	0.0038		0.9126
MYMI	*Correlation coefficient*	0.0066	0.0264	0.0027	1.0000
	P-value	0.7908	0.2884	0.9126	

Table 7. Bivariate correlations for NHL-T

		Pearson Correlation Coefficients, N = 1,616			
		Incidence	**Smoothed incidence**	**Mean distance**	**MYMI**
Incidence	*Correlation coefficient*	1.0000		-0.0225	0.0112
	P-value			0.3659	0.6528
Smoothed incidence	*Correlation coefficient*		1.0000	-0.0851	0.0197
	P-value			0.0006	0.4289
Mean distance	*Correlation coefficient*	-0.0225	-0.0851	1.0000	0.0027
	P-value	0.3659	0.0006		0.9126
MYMI	*Correlation coefficient*	0.0112	0.0197	0.0027	1.0000
	P-value	0.6528	0.4289	0.9126	

Table 7. Bivariate correlations for CLL/SLL

		Pearson Correlation Coefficients, N = 1,616			
		Incidence	**Smoothed incidence**	**Mean distance**	**MYMI**
Incidence	*Correlation coefficient*	1.0000		0.0181	0.0116
	P-value			0.4671	0.6417
Smoothed incidence	*Correlation coefficient*		1.0000	-0.0232	0.0154
	P-value			0.3523	0.5362
Mean distance	*Correlation coefficient*	0.0181	-0.0232	1.0000	0.0027
	P-value	0.4671	0.3523		0.9126
MYMI	*Correlation coefficient*	0.0116	0.0154	0.0027	1.0000
	P-value	0.6417	0.5362	0.9126	

Table 9. Poisson regression results for indirectly standardized incidence rates

Explanatory variable	NHL	NHL-B	DLBCL	FL	NHL-T	CLL/SLL
Mean distance from benzene release site						
β	-0.0036	-0.0035	-0.0031	-0.0040	-0.0016	0.0012
Standard error	0.0001	0.0001	0.0001	0.0002	0.0002	0.0005
Wald χ^2 statistic	963.81	901.68	3,142.61	404.25	60.89	6.69
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0010	0.0097
Model fit						
Deviance, df = 1,614	29624.29	36050.70	231021.13	36161.43	28700.25	5974.15
Deviance/df	18.35	22.33	143.14	22.40	17.78	3.70
Incidence rate ratios						
1.0 miles from site	0.9964	0.9965	0.9968	0.9960	0.9984	1.0012
(95% confidence interval)	(0.9962-0.9966)	(0.9962-0.9967)	(0.9968-0.9971)	(0.9956-0.9964)	(0.9980-0.9988)	(1.003-1.0022)
99.5 miles from site	0.6981	0.7025	0.7379	0.6716	0.8502	1.1305
(95% confidence interval)	(0.6824-0.7141)	(0.6865-0.7189)	(0.7301-0.7458)	(0.6461-0.6982)	(0.8163-0.8856)	(1.0302-1.2406)

Table 10. Poisson regression results for SEB-smoothed indirectly standardized rates

Explanatory variable	NHL	NHL-B	DLBCL	FL	NHL-T	CLL/SLL
Mean distance from benzene release site						
β	-0.0043	-0.0031	-0.0035	-0.0020	-0.0032	-0.0010
Standard error	0.0001	0.0001	0.0001	0.0003	0.0003	0.0008
Wald χ^2 statistic	1,098.42	548.28	2,657.07	55.72	100.97	1.50
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.2199
Model fit						
Deviance, df = 1,614	13,698.61	15,315.59	94,512.00	8,964.21	8,628.04	1,670.76
Deviance/df	8.49	9.49	58.56	5.55	5.35	1.04
Incidence rate ratios						
1.0 miles from site	0.9957	0.9969	0.9965	0.9980	0.9968	0.9990
(95% confidence interval)	(0.9955-0.9960)	(0.9967-0.9972)	(0.9964-0.9967)	(0.9975-0.9985)	(0.9961-0.9974)	(0.9973-1.0006)
99.5 miles from site	0.6534	0.7367	0.7075	0.8203	0.7237	0.9033
(95% confidence interval)	(0.6371-0.670	(0.7181-0.755	(0.6982-0.716	(0.7787-0.864	(0.6795-0.770	(0.7678-1.0627)

Table 11. Global measures of spatial autocorrelation

		NHL	NHL-B	DLBCL	FL	NHL-T	CLL/SLL
SIR	*Moran's I*	0.2009	0.0628	0.0266	-0.0011	0.0430	0.0031
	p-value	0.0010	0.0010	0.0360	0.5150	0.0020	0.3590
SEB-smoothed SIR	*Moran's I*	0.8317	0.6377	0.6478	0.5938	0.5879	0.6002
	p-value	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010

Table 12. Lawson-Waller score test for focal clustering of NHL near benzene release sites

	Incidence		SEB-smoothed incidence	
	Score	*p-value*	*Score*	*p-value*
Site 1	13.1831	<0.0001	12.8390	<0.0001
Site 2	18.0533	<0.0001	19.4879	<0.0001
Site 3	36.3879	<0.0001	29.5851	<0.0001
Site 4	14.9471	<0.0001	15.3119	<0.0001
Site 5	15.6043	<0.0001	15.9281	<0.0001
Site 6	-23.6137	1.0000	-26.3927	1.0000
Site 7	15.8019	<0.0001	16.1439	<0.0001
Site 8	-3.8011	0.9999	-3.3802	0.9996
Site 9	-3.4838	0. 9999	-4.5192	1.0000
Site 10	-6.7270	1.0000	-7.1492	1.0000
Site 11	26.9324	<0.0001	15.5133	<0.0001
Site 12	20.8564	<0.0001	13.7356	<0.0001
Site 13	14.9068	<0.0001	15.3023	<0.0001
Site 14	15.6043	<0.0001	15.9281	<0.0001
Site 15	15.0695	<0.0001	15.1875	<0.0001
Site 16	9.2829	<0.0001	1.2391	0.1077
Site 17	29.9130	<0.0001	25.8678	<0.0001
Site 18	37.9698	<0.0001	31.8159	<0.0001
Site 19	18.4694	<0.0001	20.7593	<0.0001

FIGURES

Figure 1. Distribution of mean distance from benzene release site

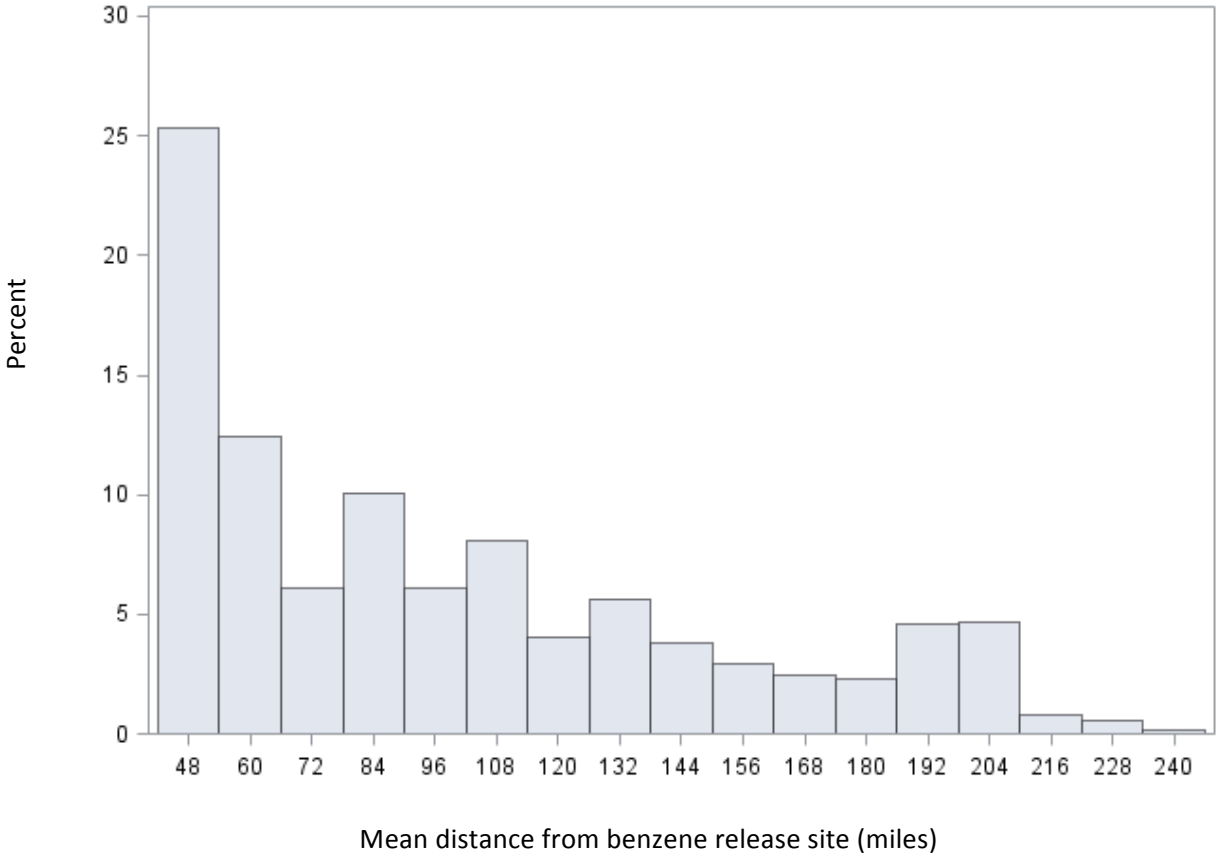


Figure 2. SIR map for NHL

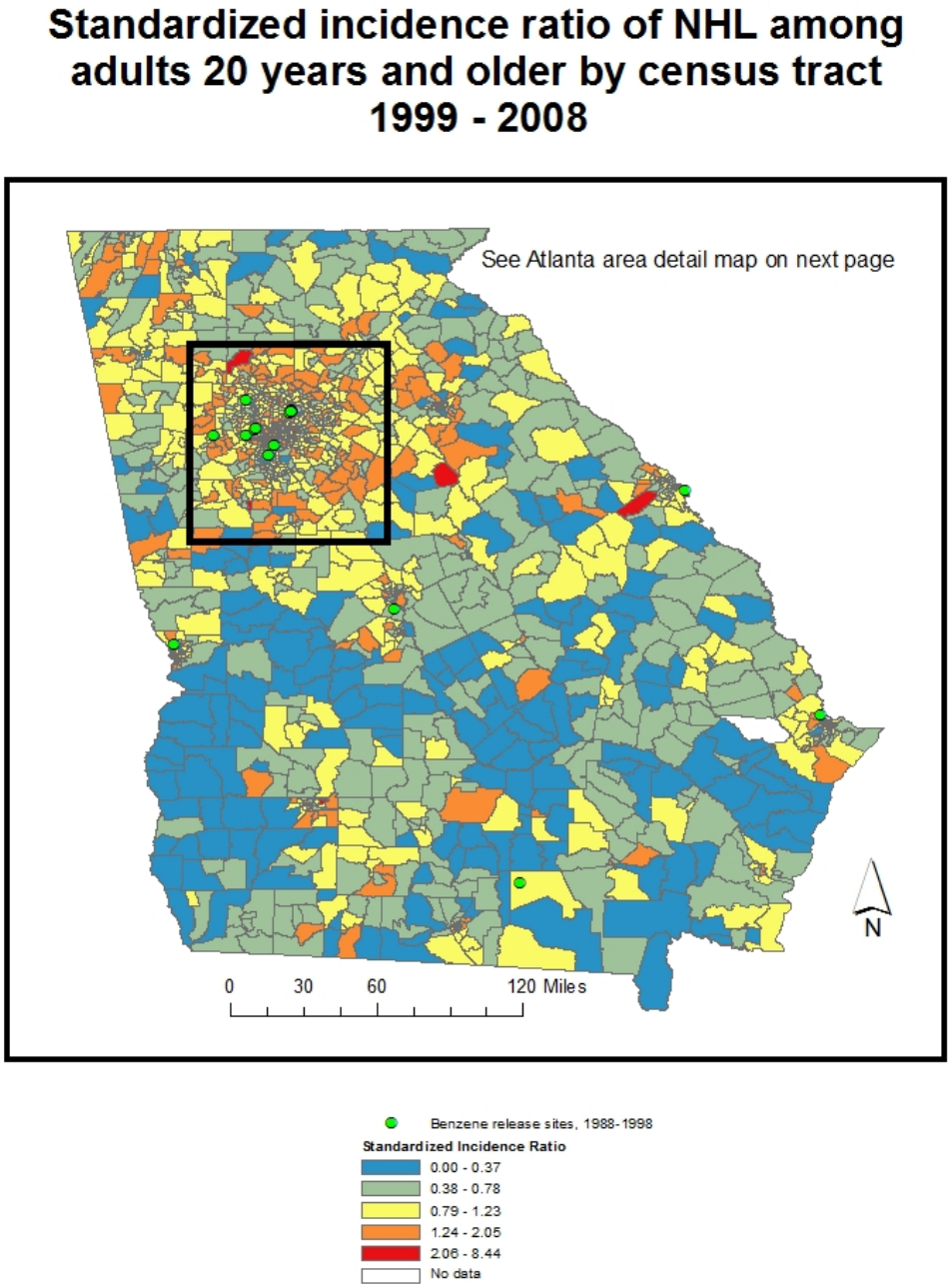


Figure 3. SIR map for NHL, metro-Atlanta area

Figure 4. SEB-smoothed SIR map for NHL

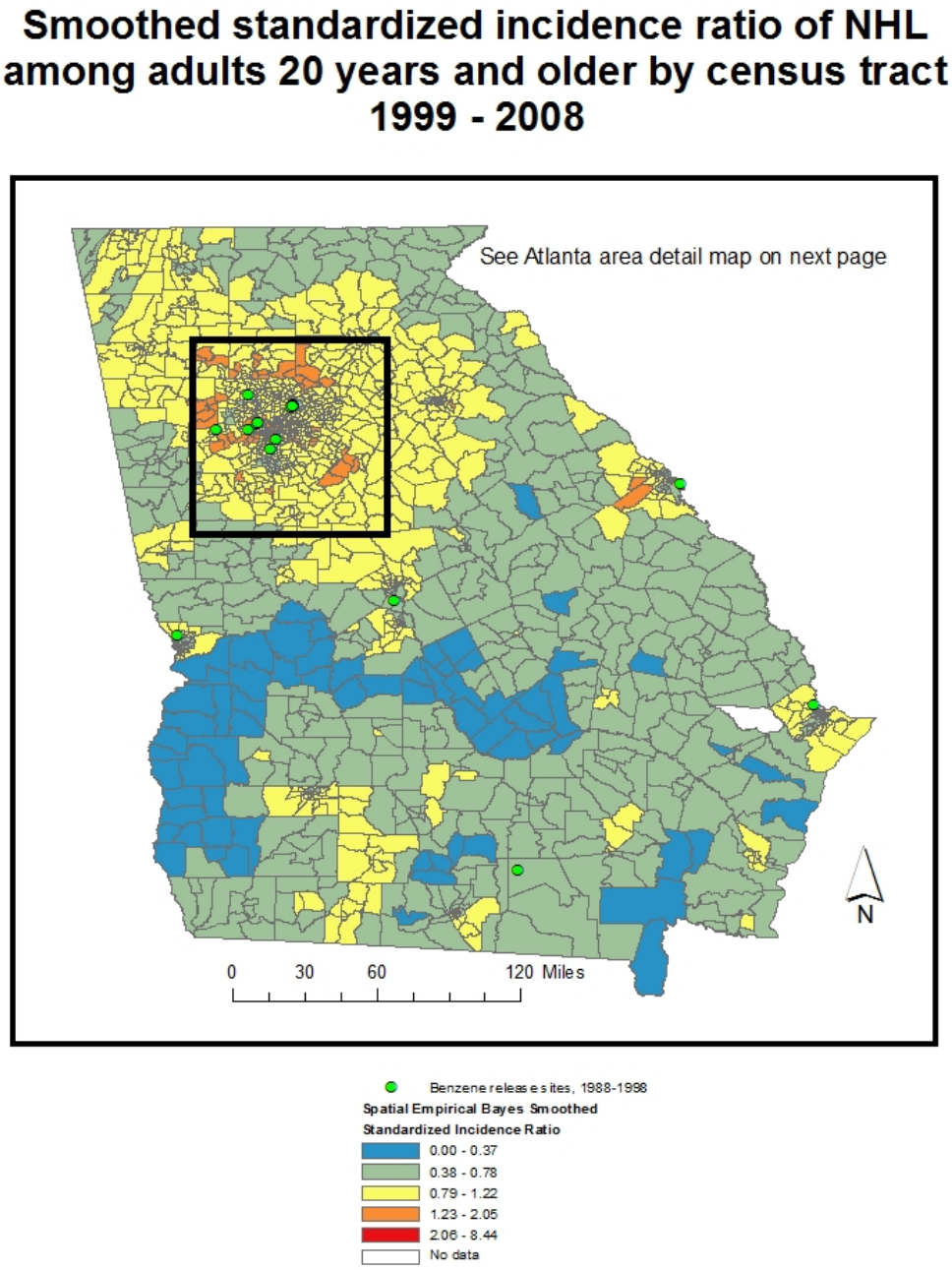


Figure 5. SEB-smoothed SIR map for NHL, metro-Atlanta area

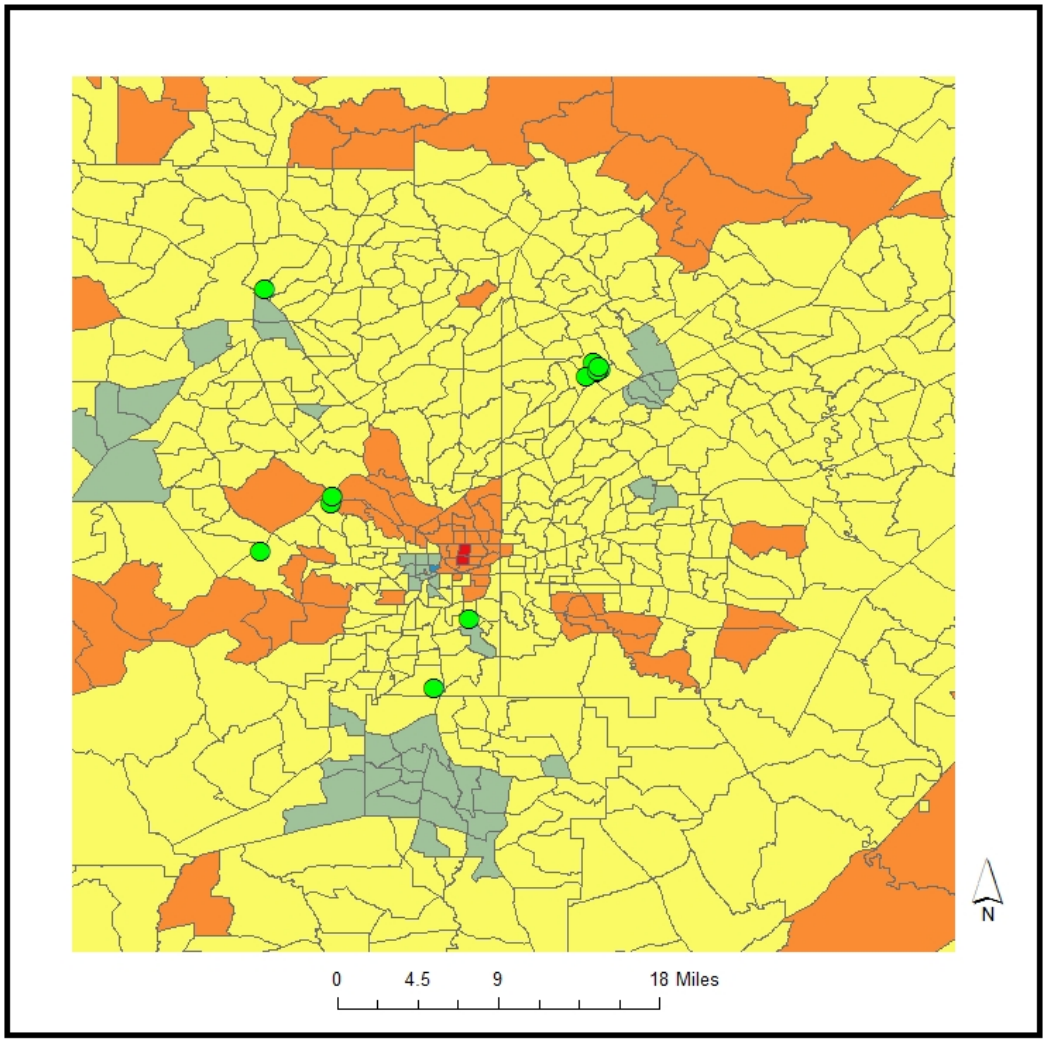


Figure 6. SIR map for NHL-B

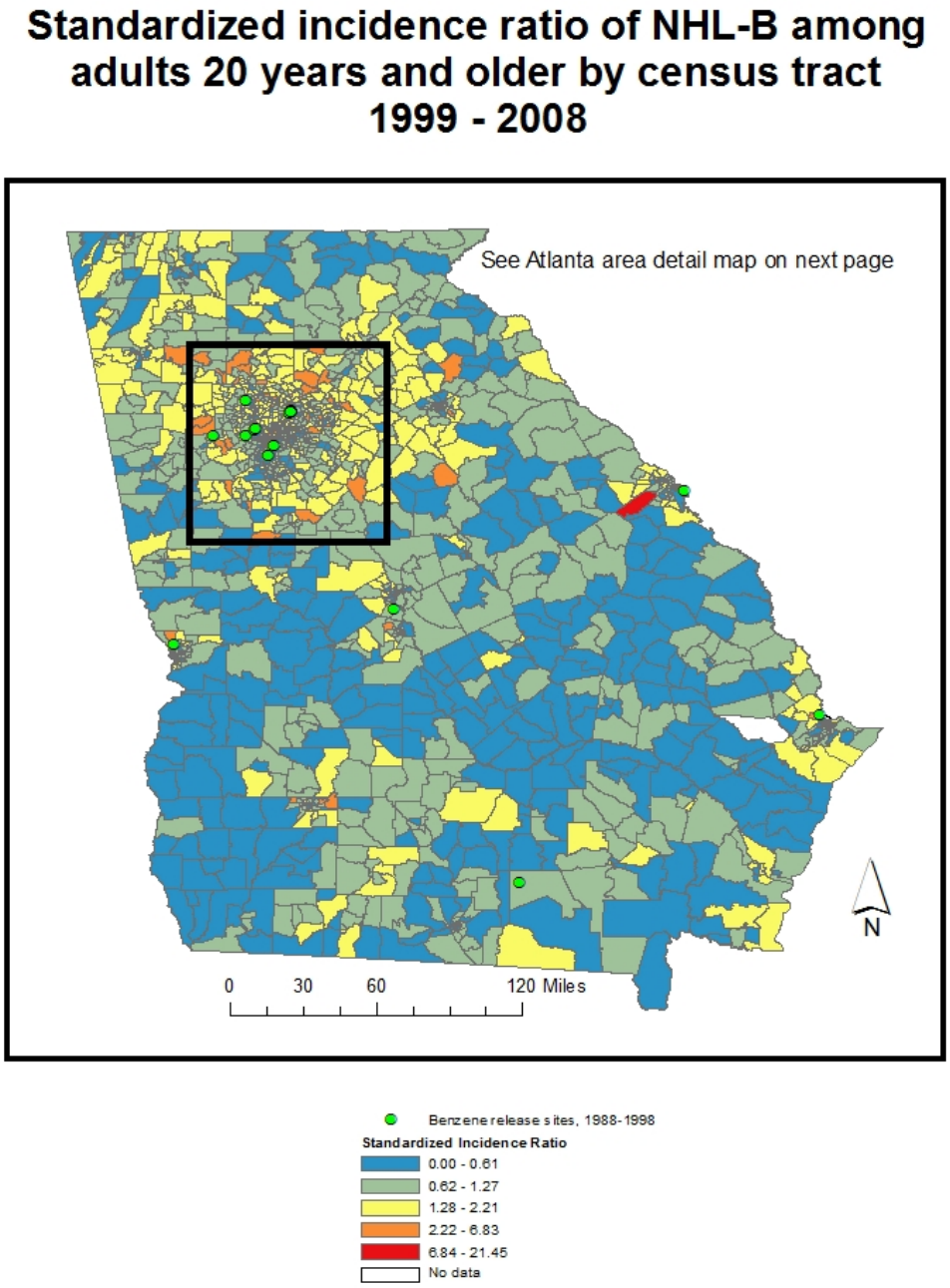


Figure 7. SIR map for NHL-B, metro-Atlanta area

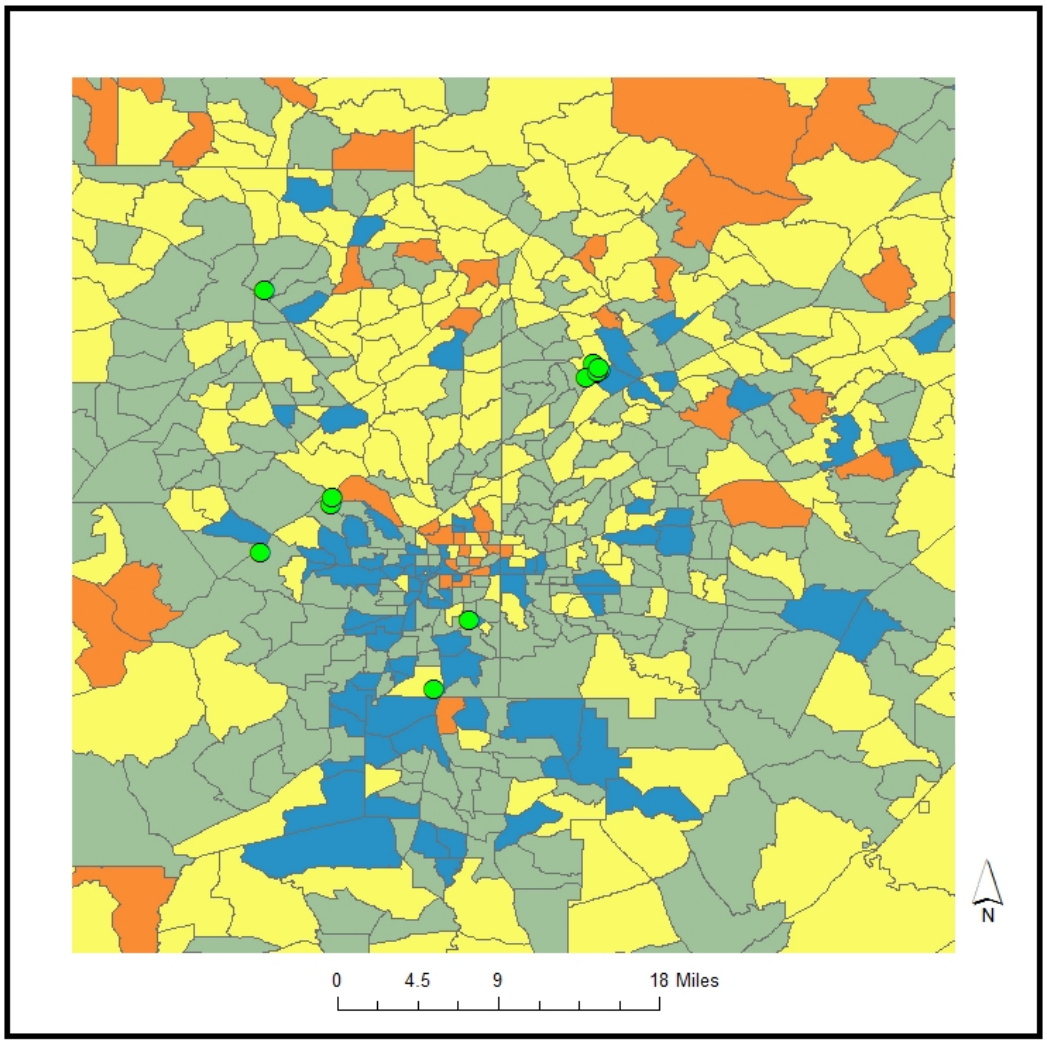


Figure 8. SEB-smoothed SIR map for NHL-B

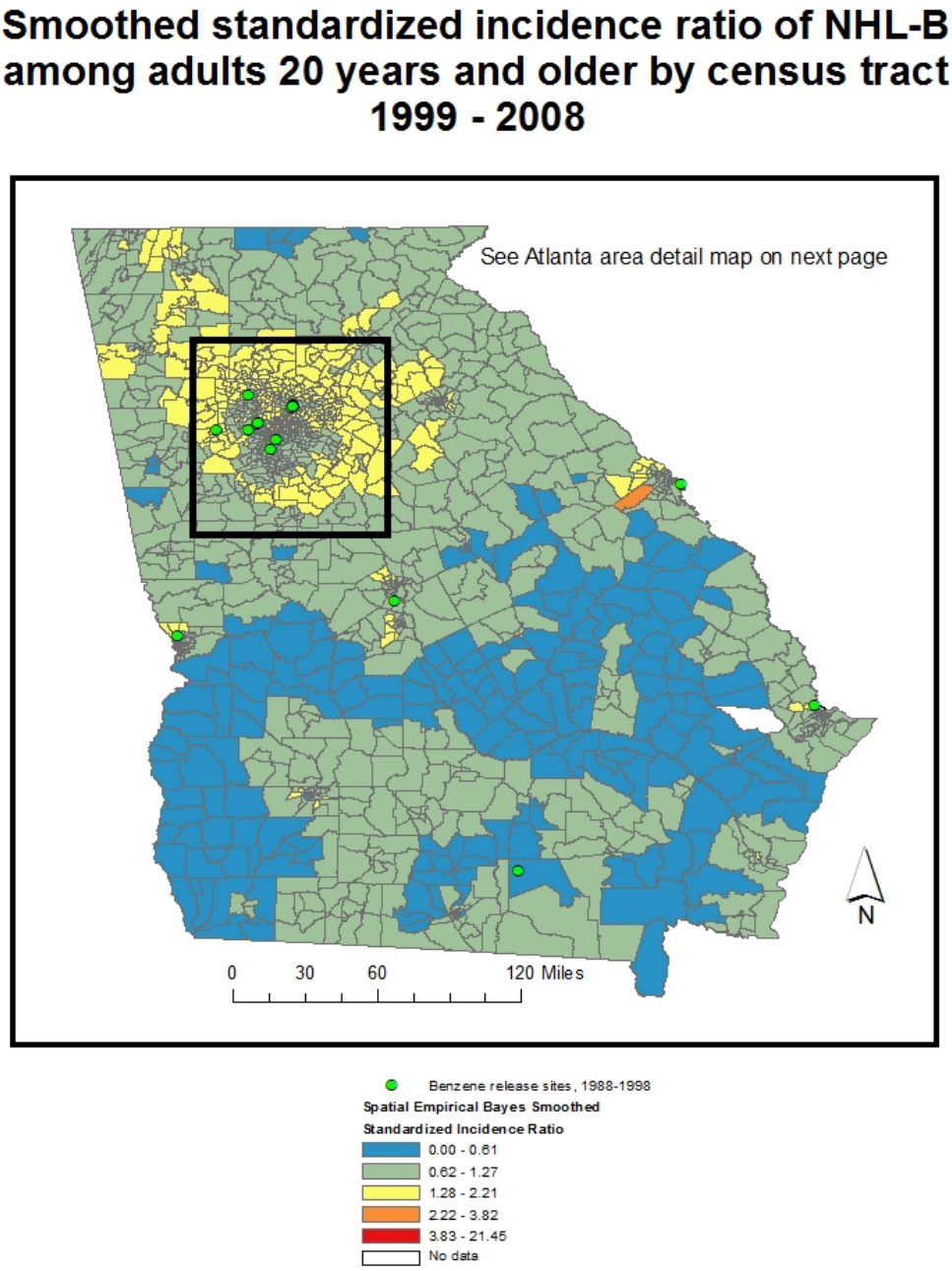


Figure 9. SEB-smoothed SIR map for NHL-B, metro-Atlanta area

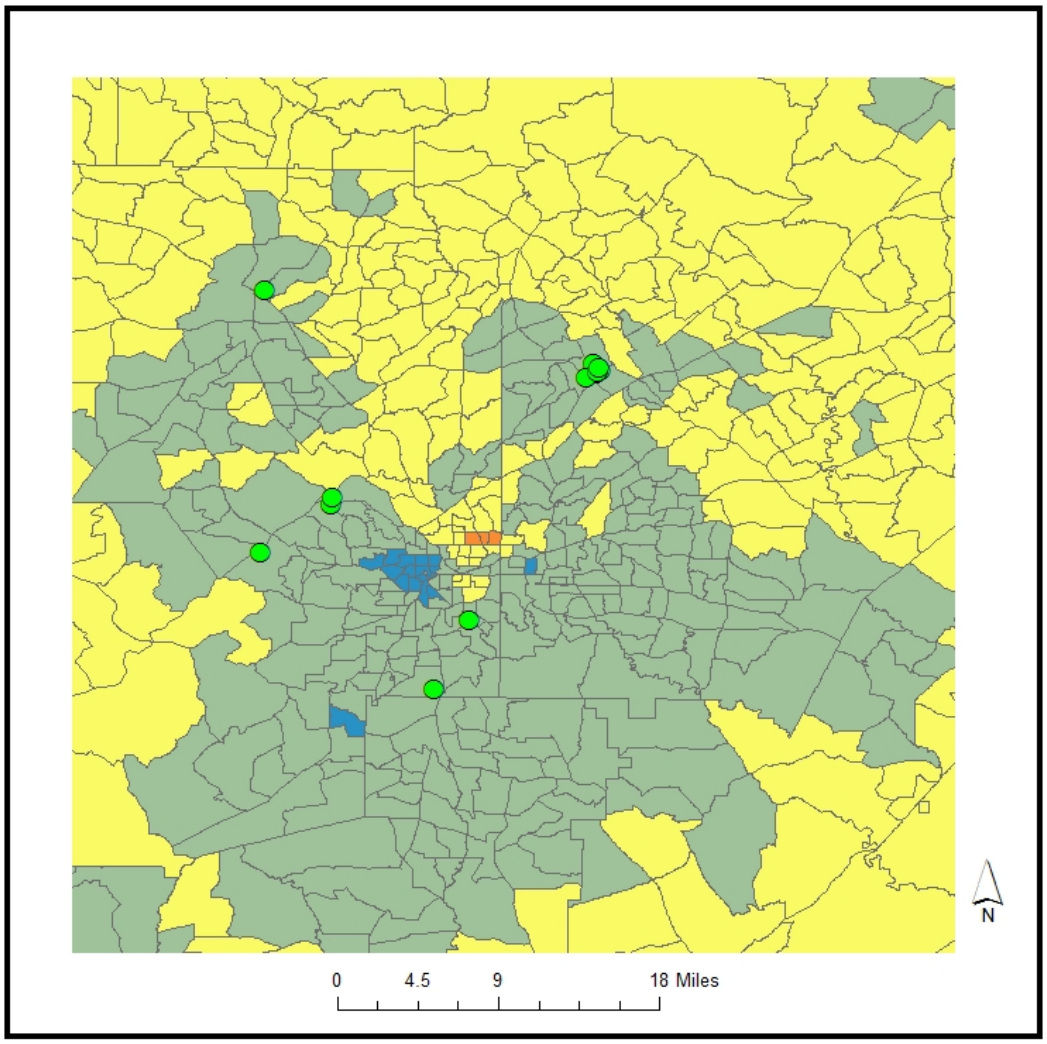


Figure 10. SIR map of DLBCL

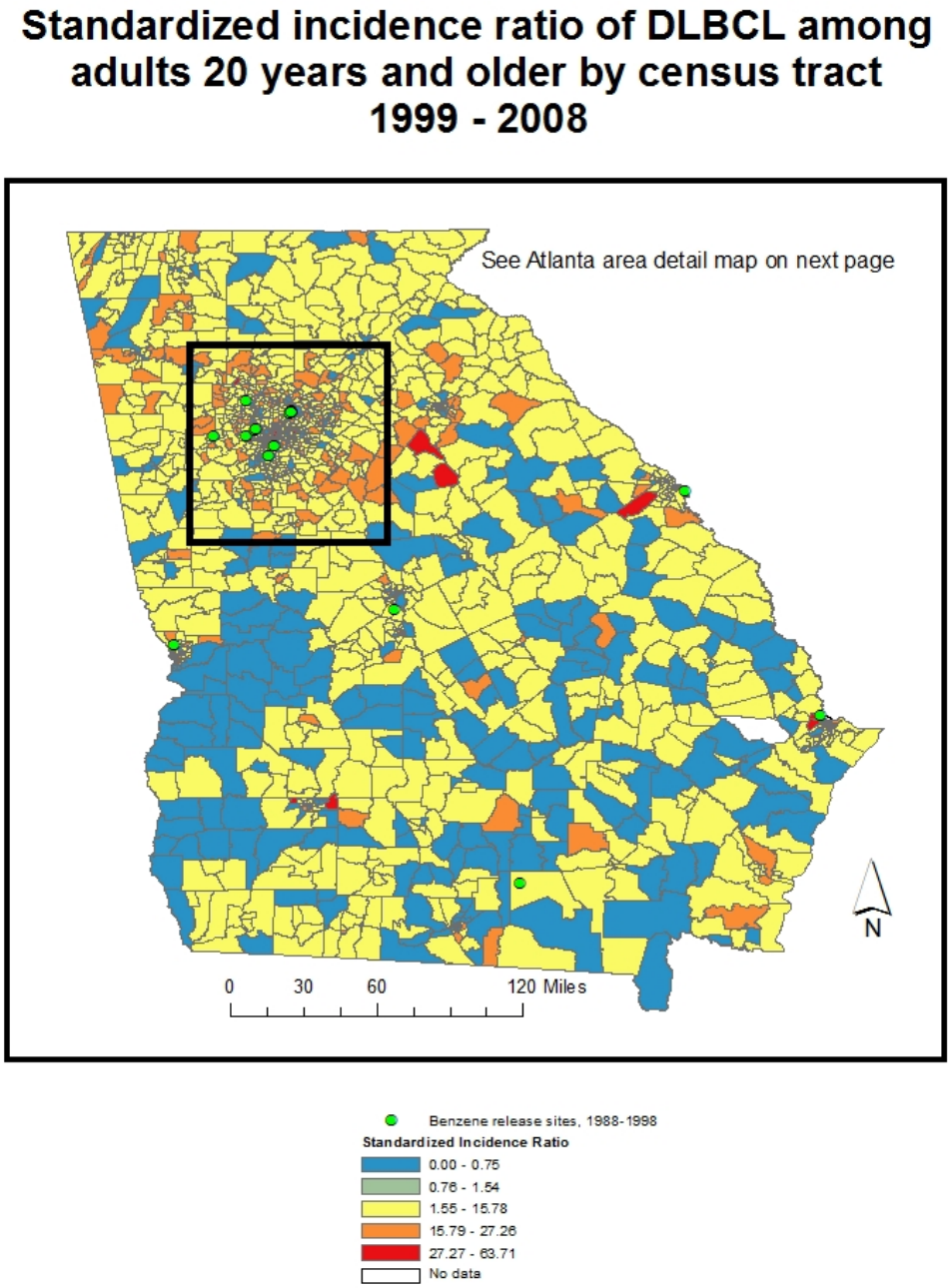


Figure 11. SIR map of DLBCL, metro-Atlanta area

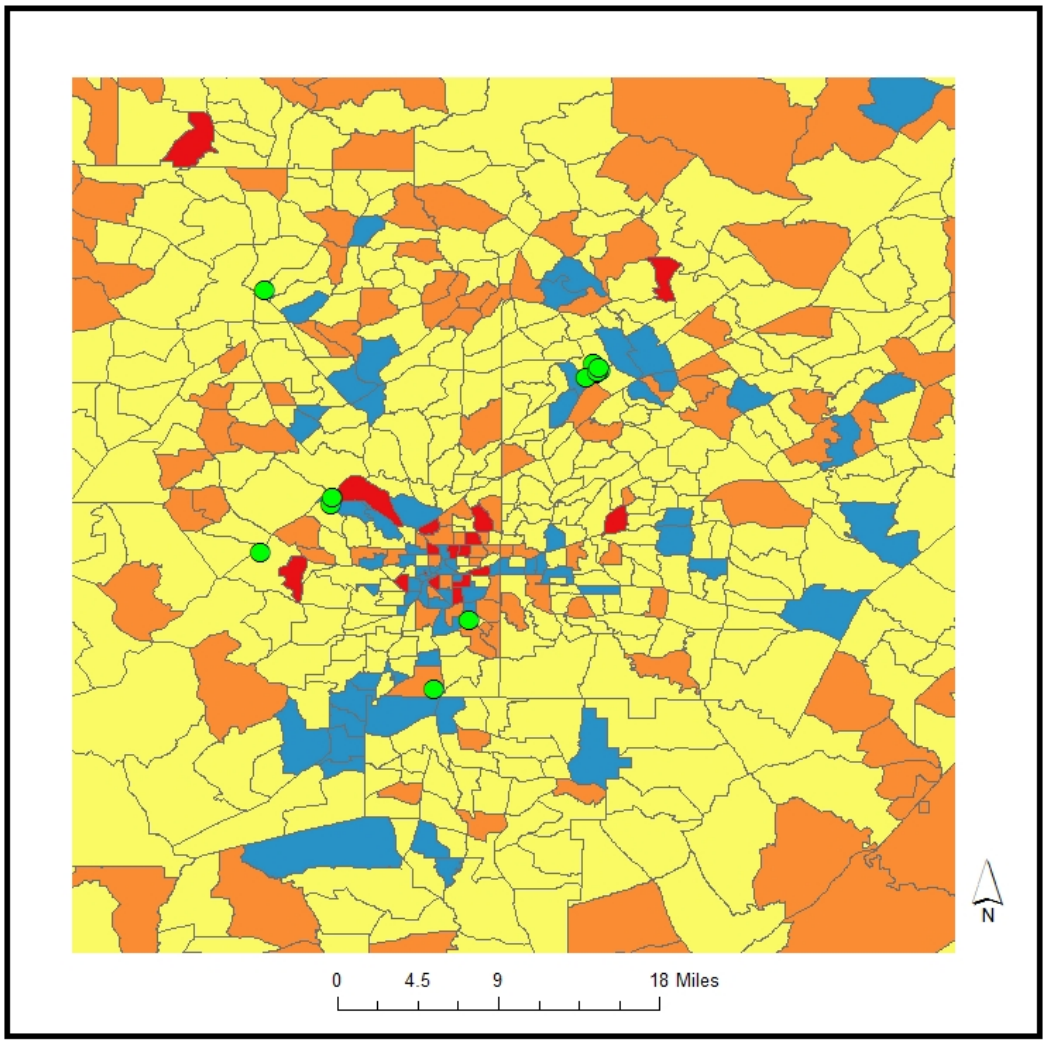


Figure 12. SEB-smoothed SIR map of DLBCL

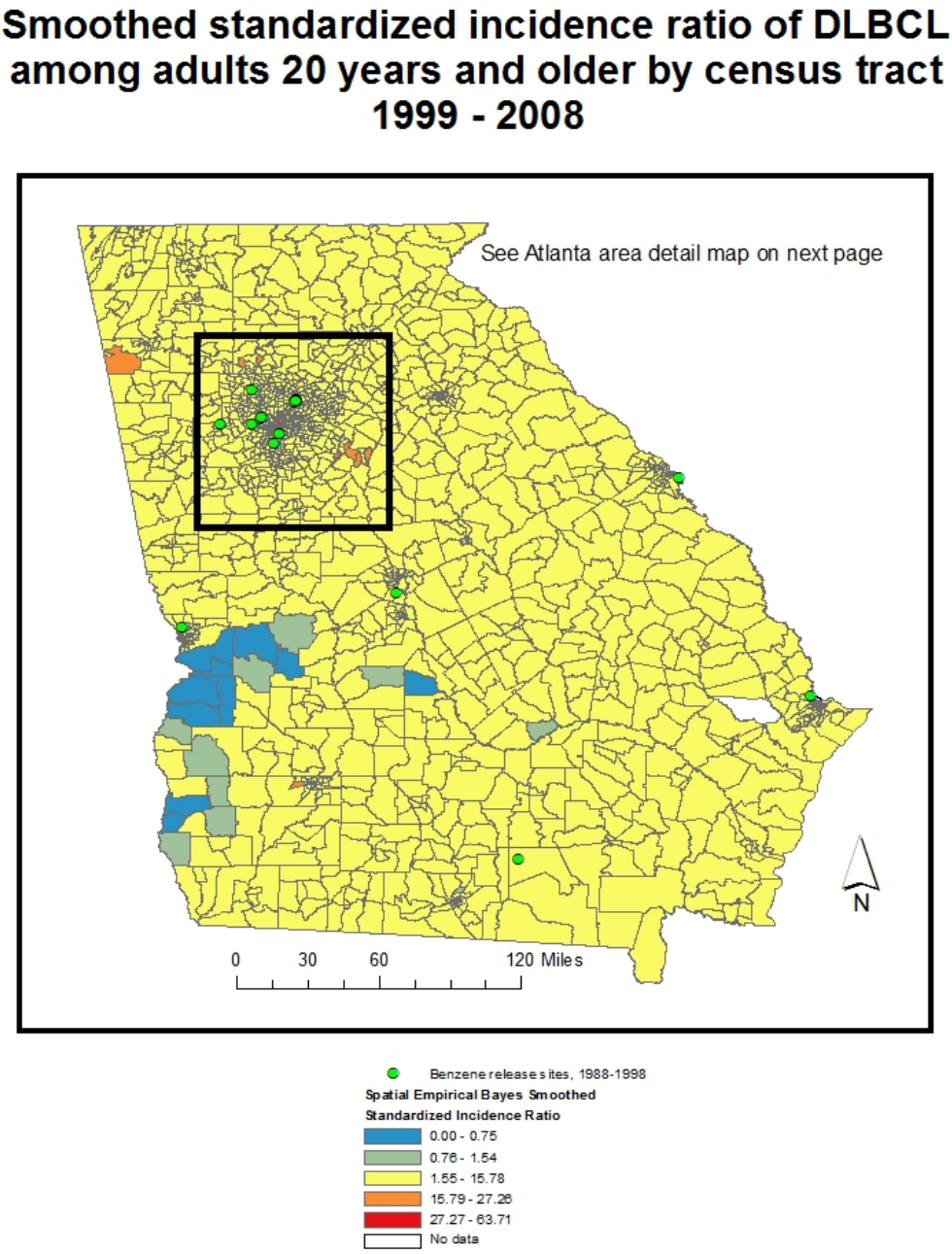


Figure 13. SEB-smoothed SIR map of DLBCL, metro-Atlanta area

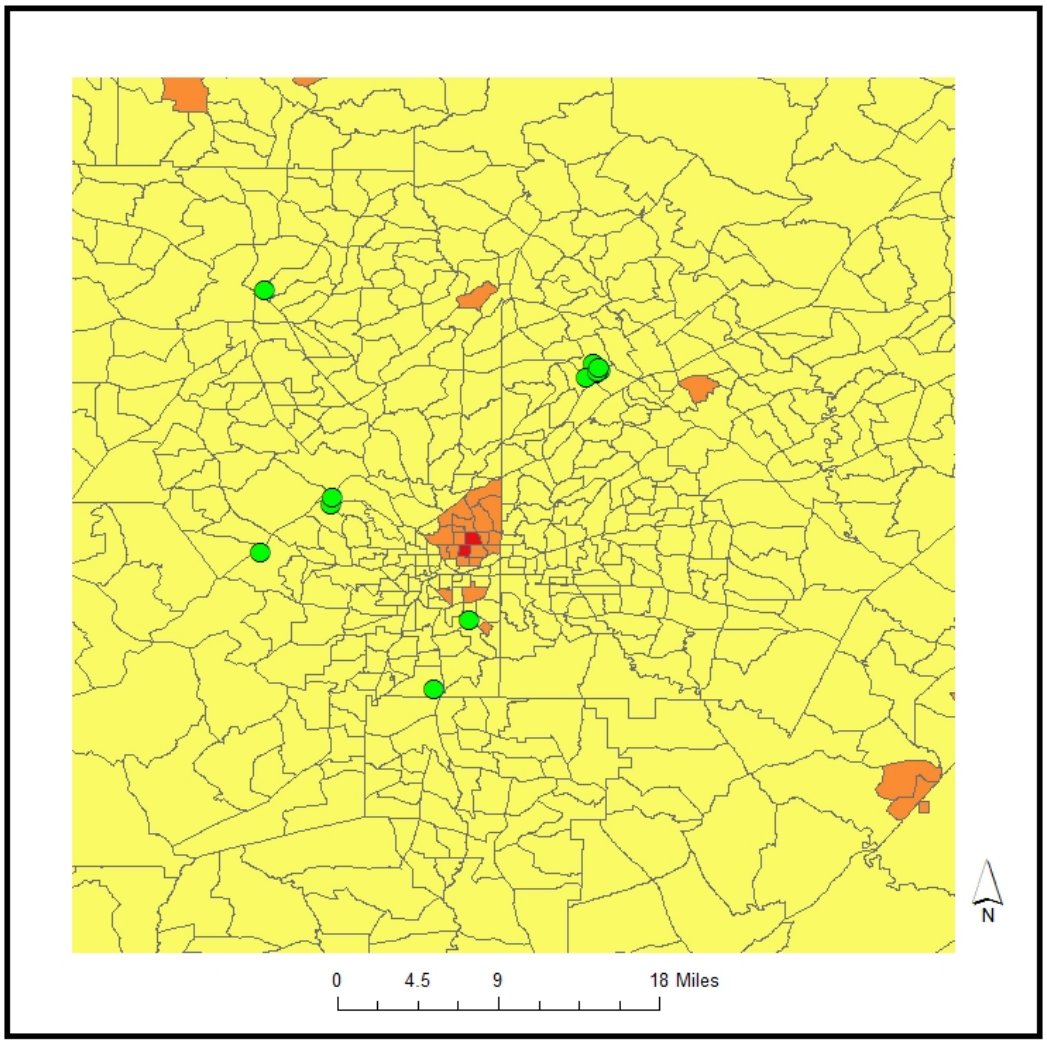


Figure 14. SIR map of FL

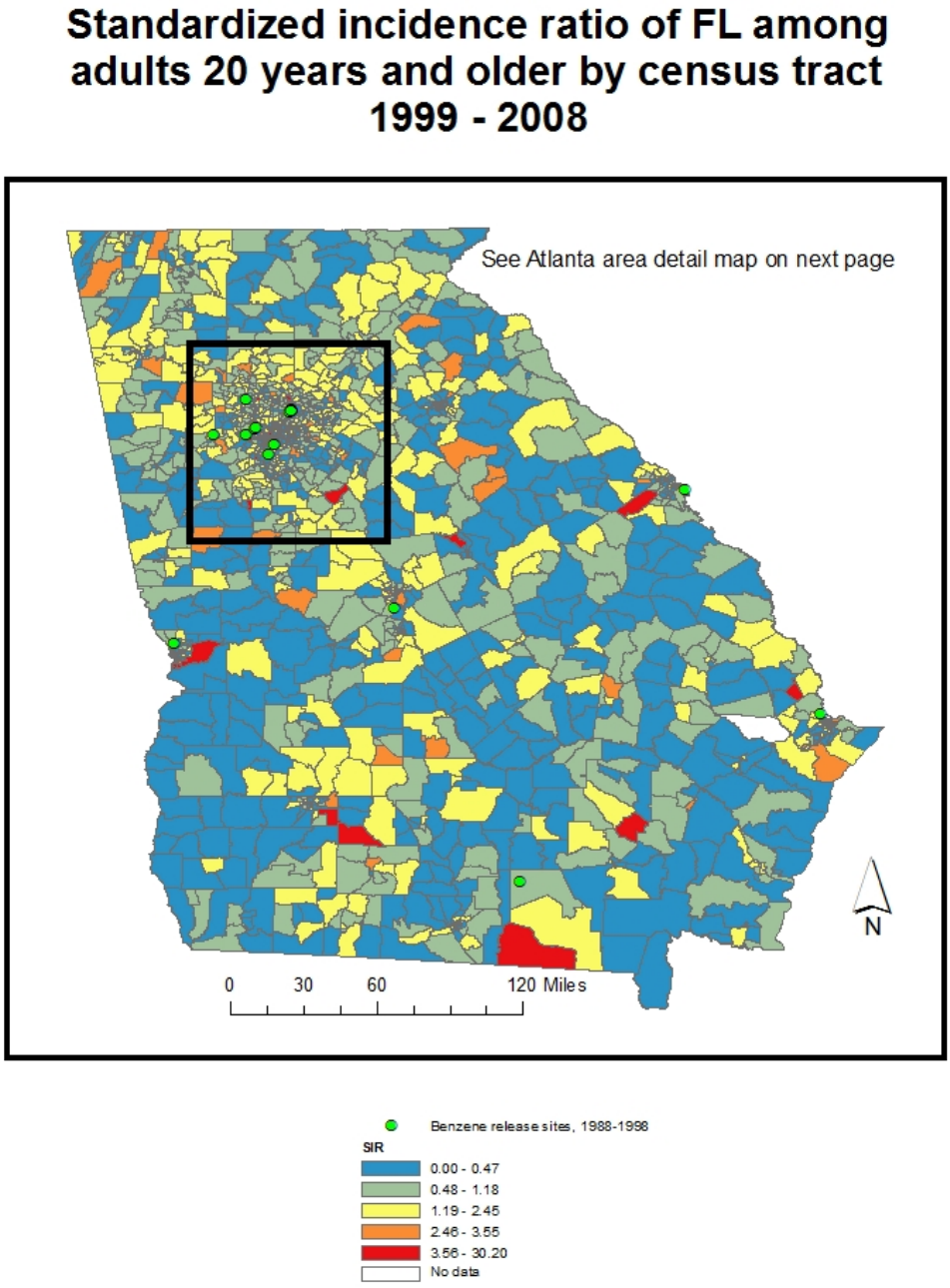


Figure 15. SIR map of FL, metro-Atlanta area

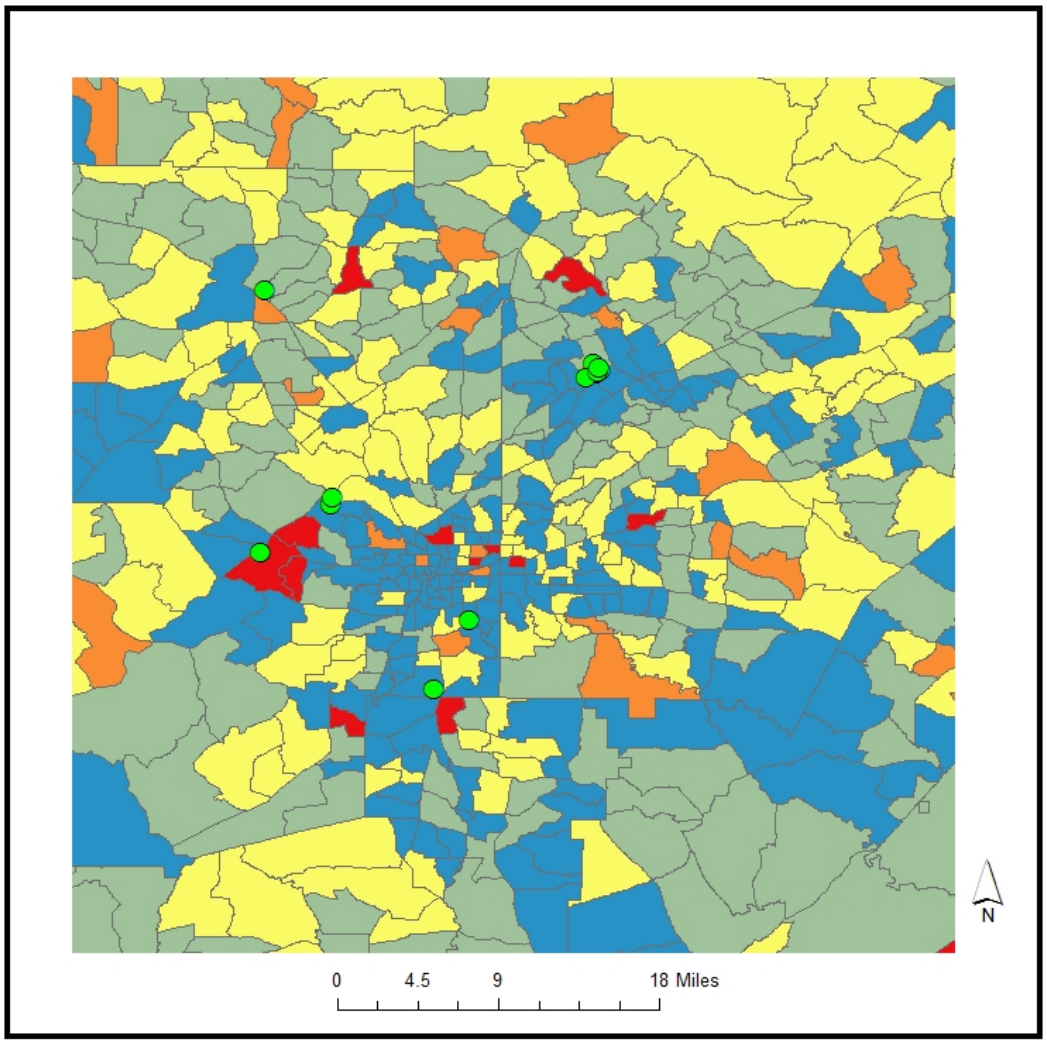


Figure 16. SEB-smoothed SIR map of FL

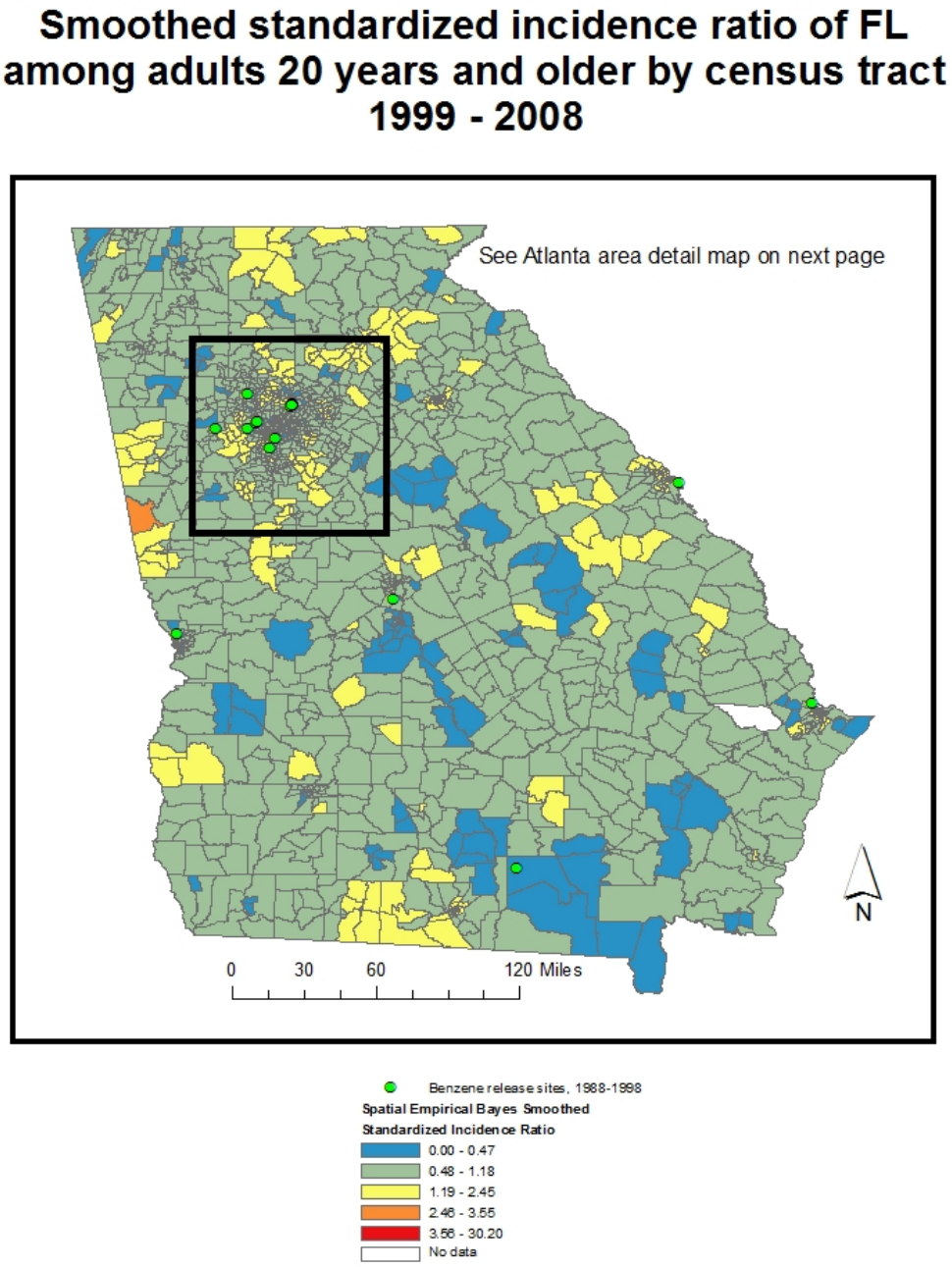


Figure 17. SEB-smoothed SIR map of FL, metro-Atlanta area

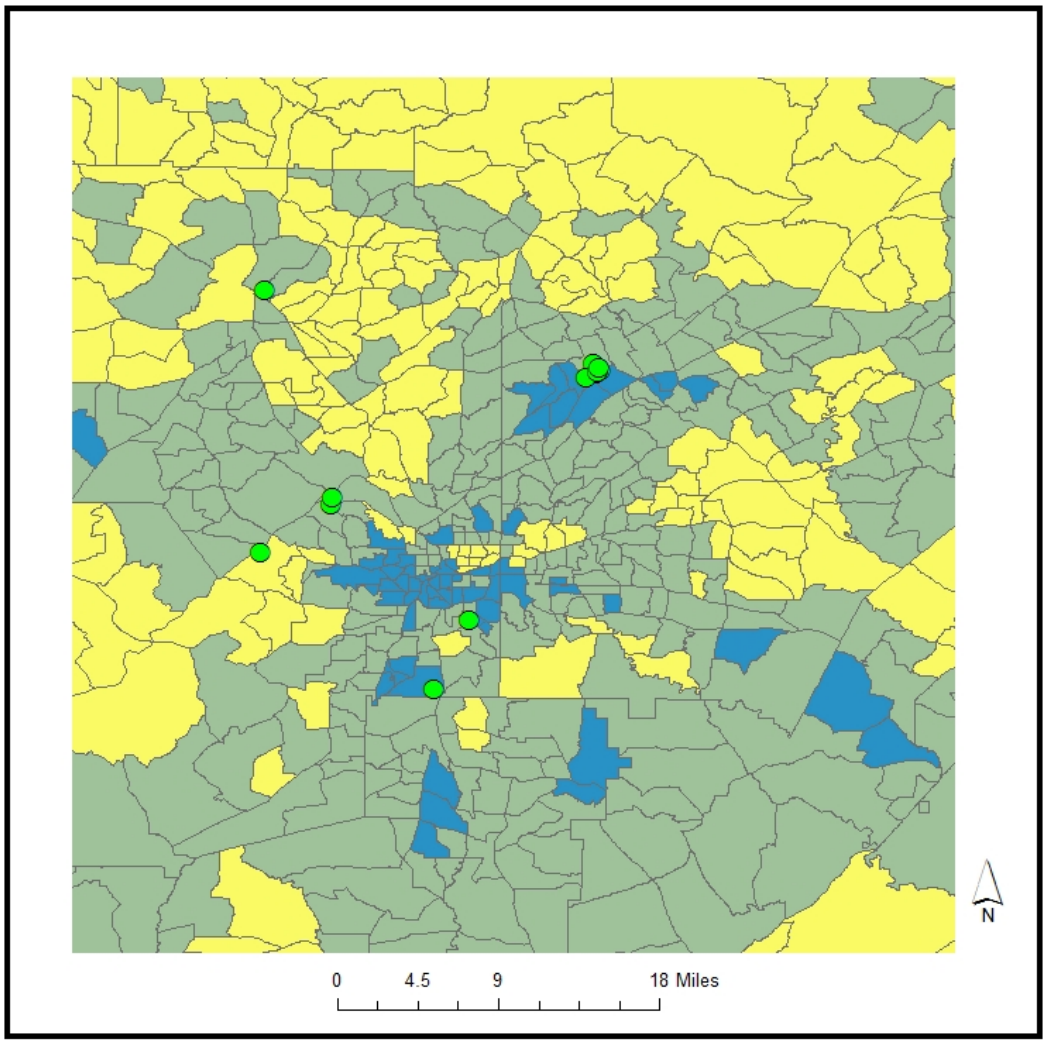


Figure 18. SIR map of NHL-T

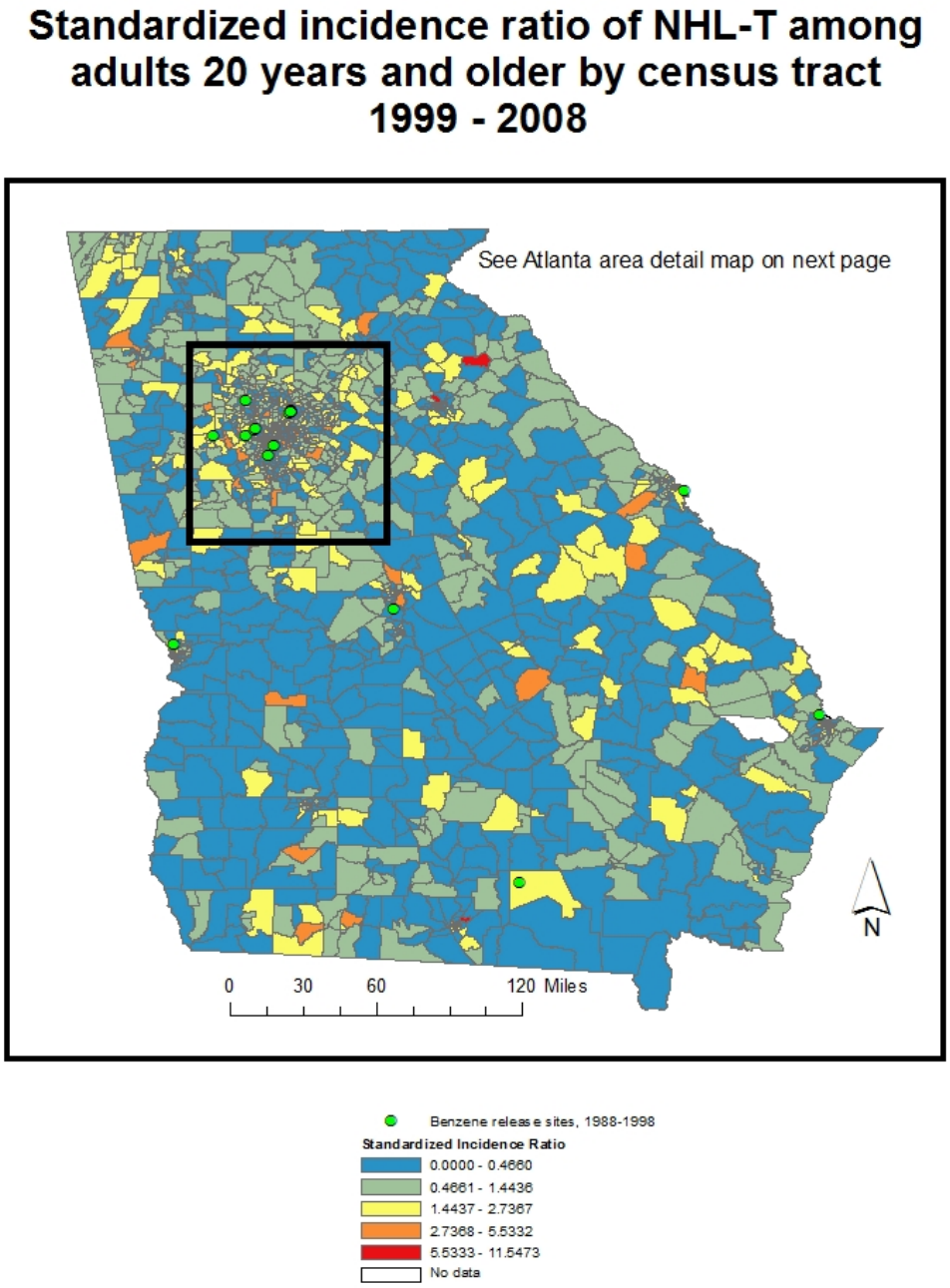


Figure 19. SIR map of NHL-T, metro-Atlanta area

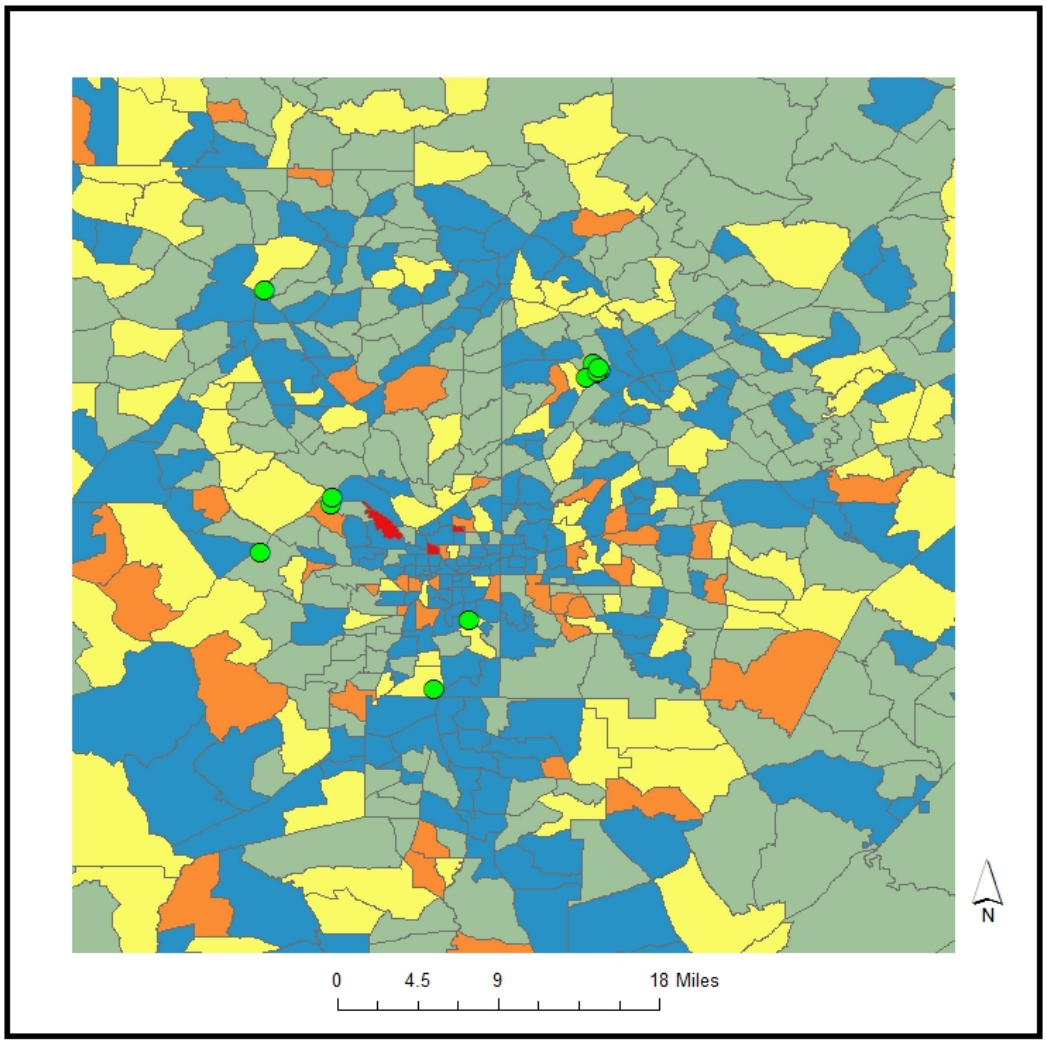


Figure 20. SEB-smoothed SIR map of NHL-T

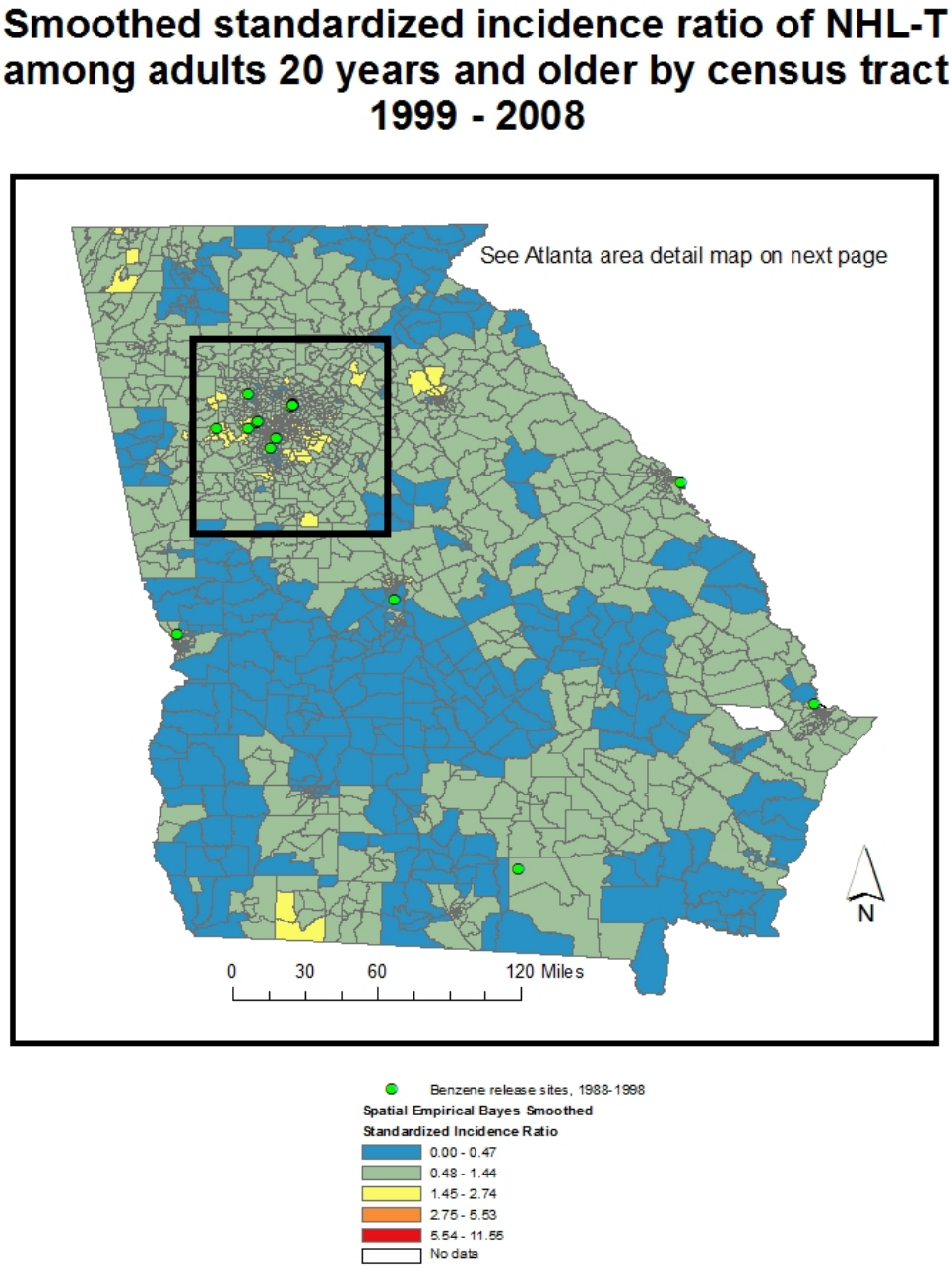


Figure 21. SEB-smoothed SIR map of NHL-T, metro-Atlanta area

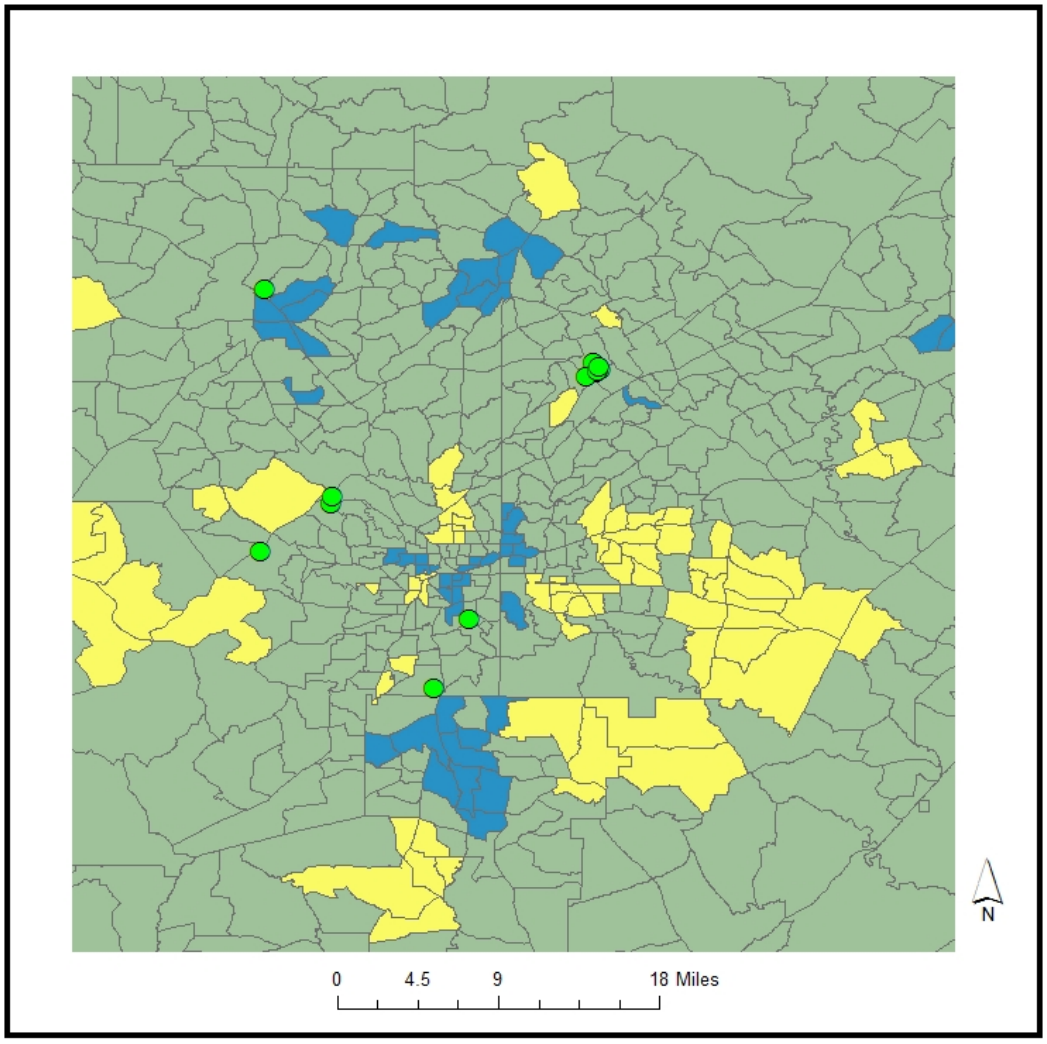


Figure 22. SIR map of CLL/SLL

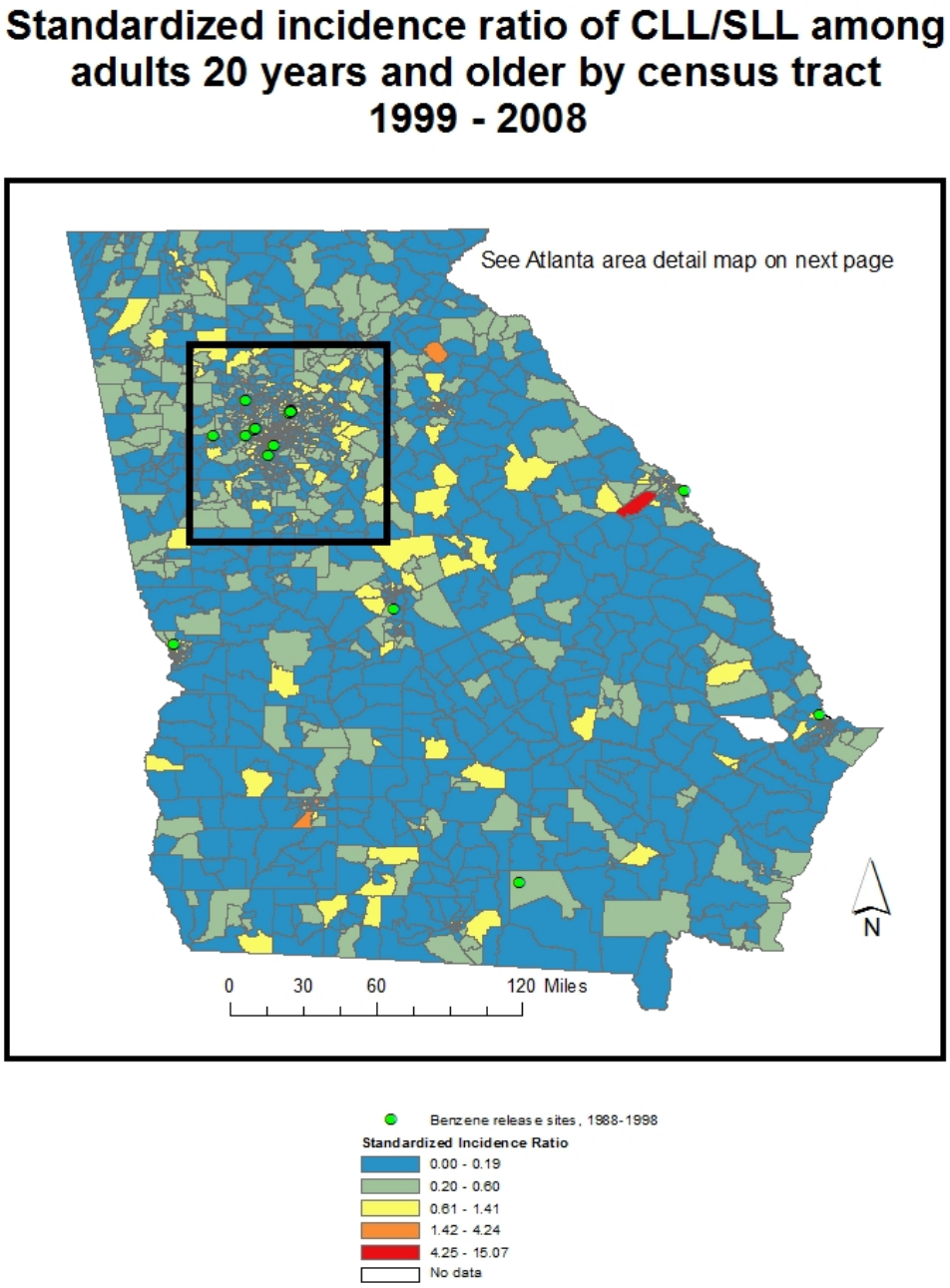


Figure 23. SIR map of CLL/SLL, metro-Atlanta area

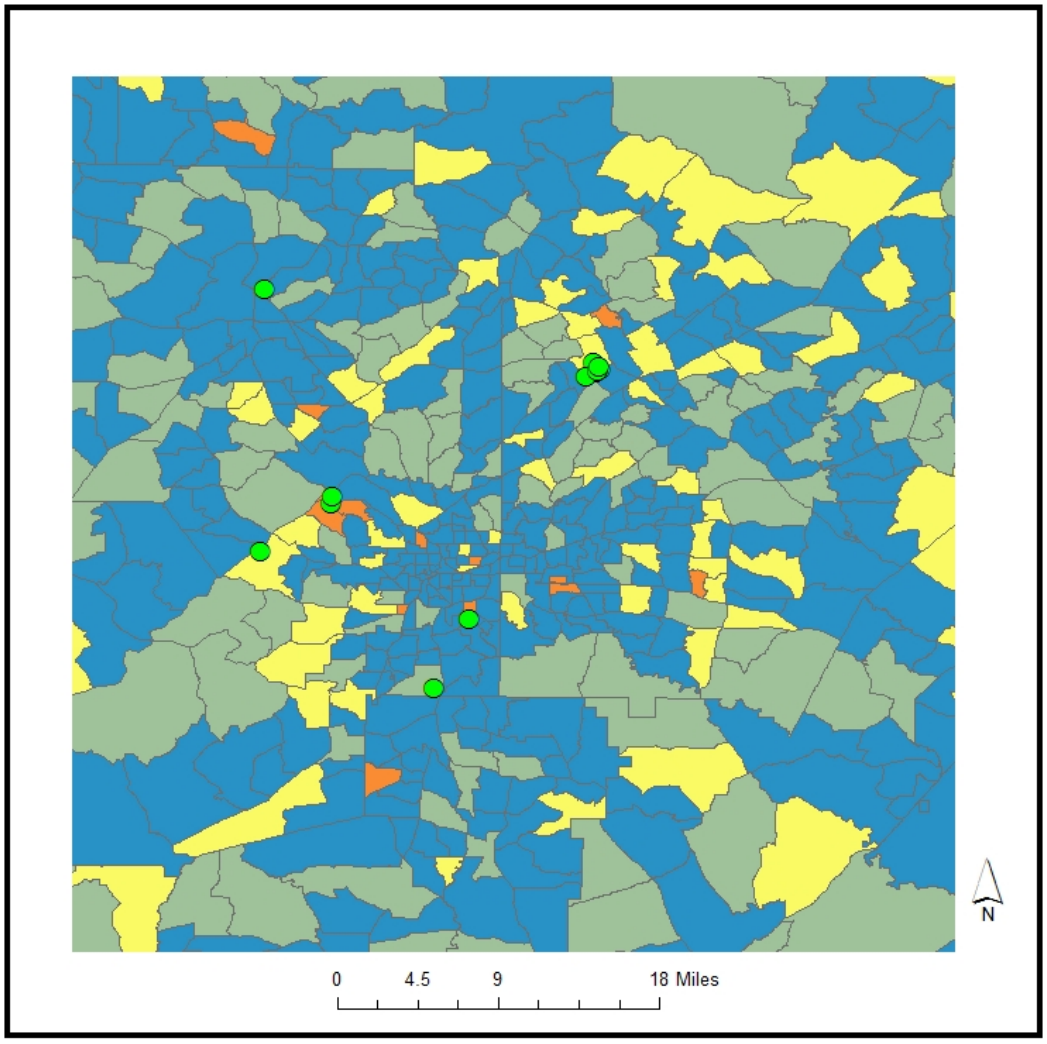


Figure 24. SEB-smoothed SIR map of CLL/SLL

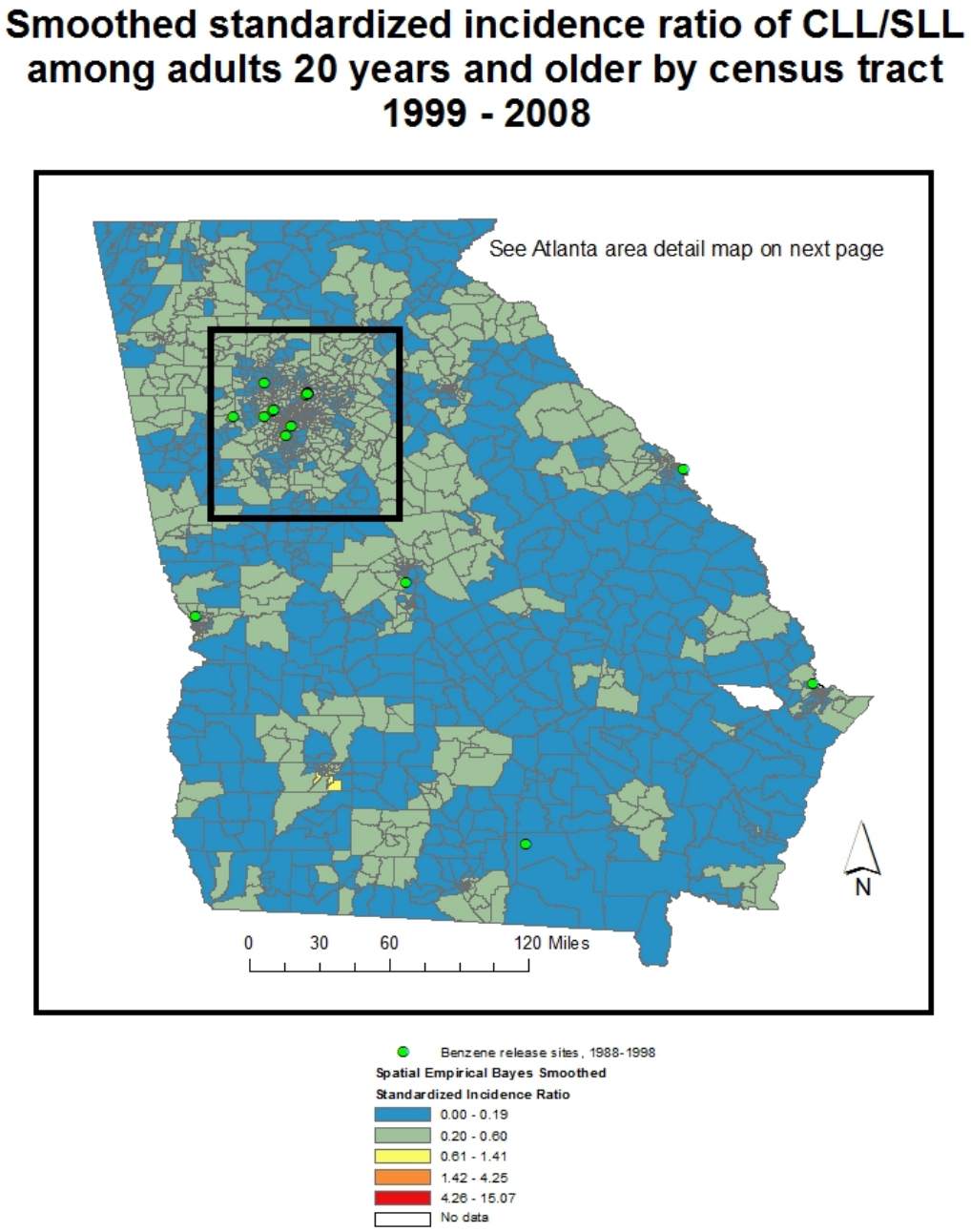


Figure 25. SEB-smoothed SIR map of CLL/SLL, metro-Atlanta area

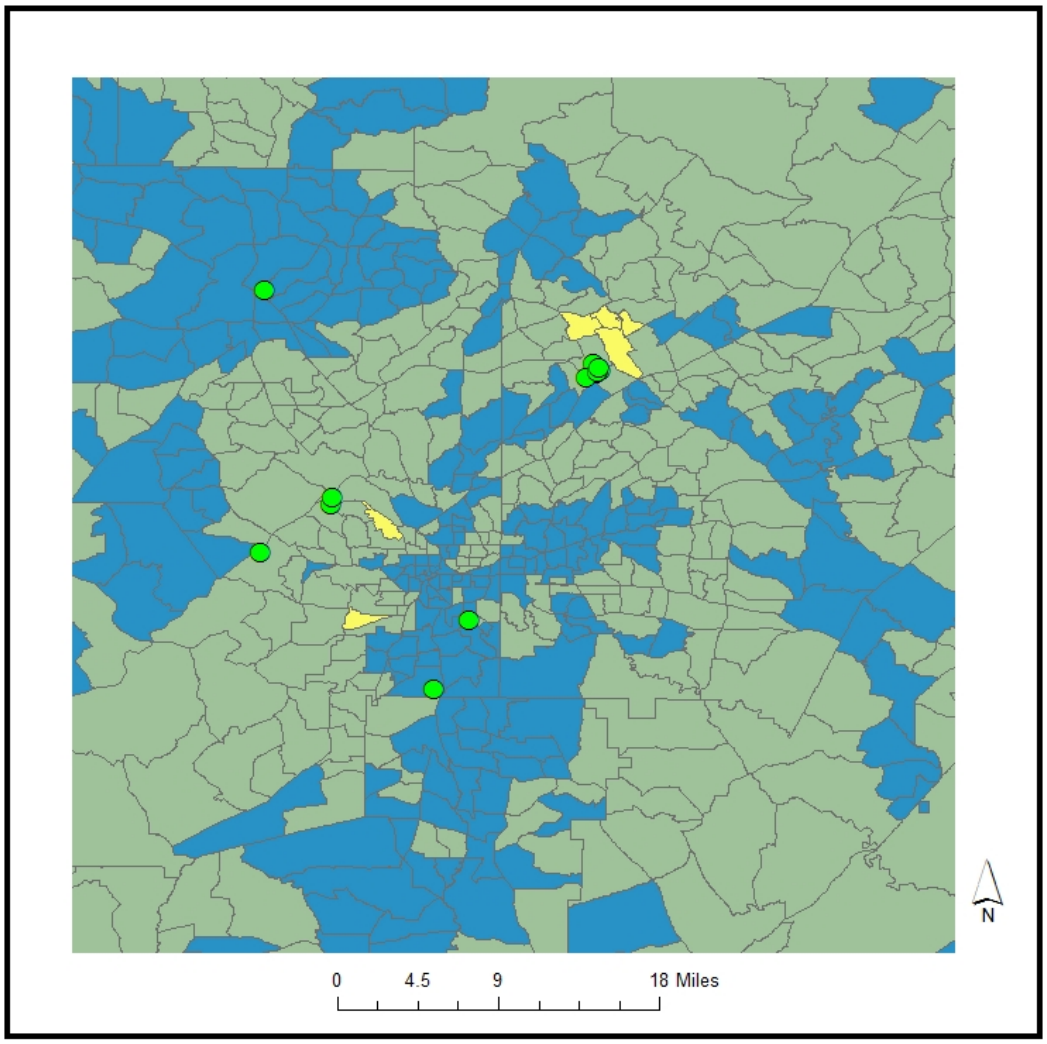


Figure 26. Cluster map of NHL SIRs

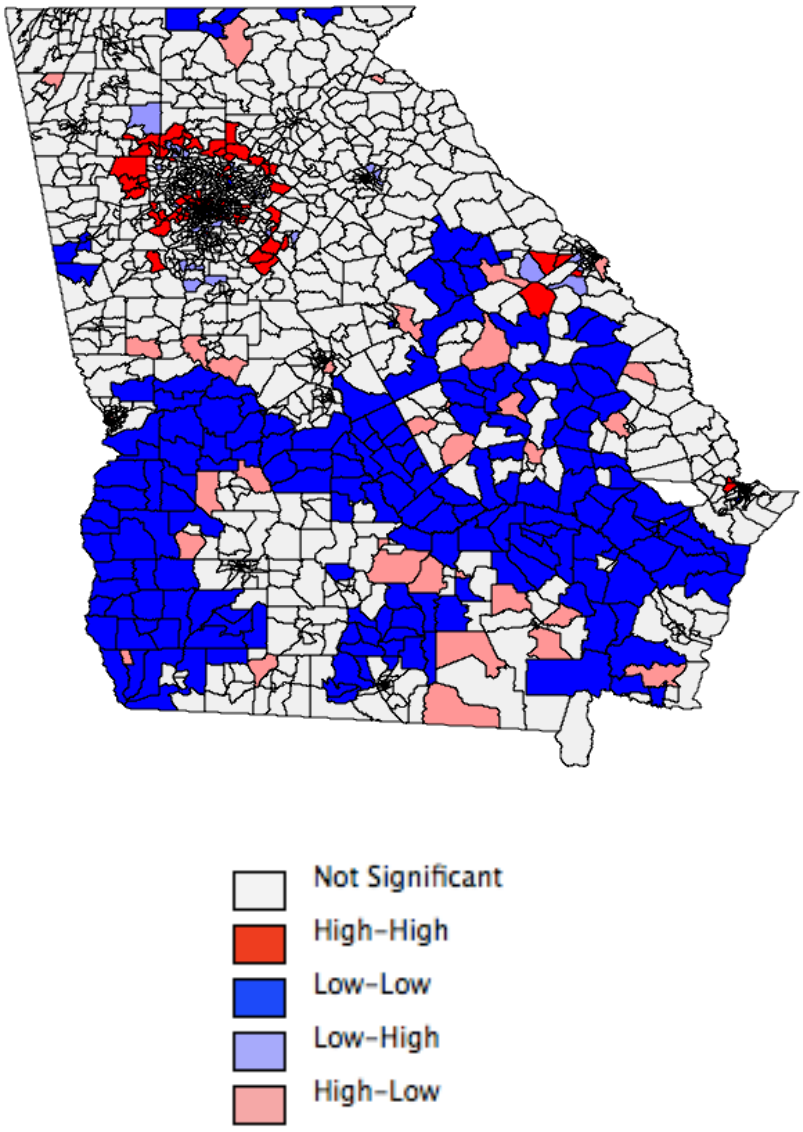


Figure 27. Cluster map of NHL SEB-smoothed SIRs

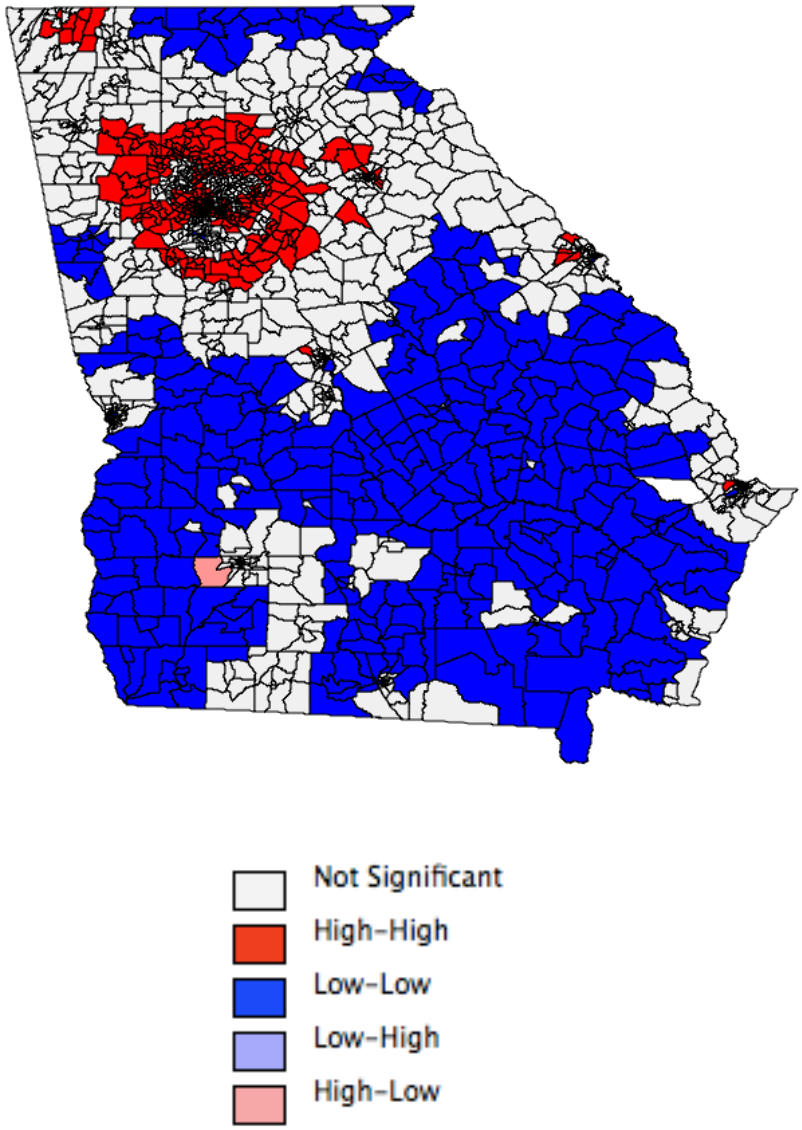


Figure 28. Cluster map of NHL-B SIRs

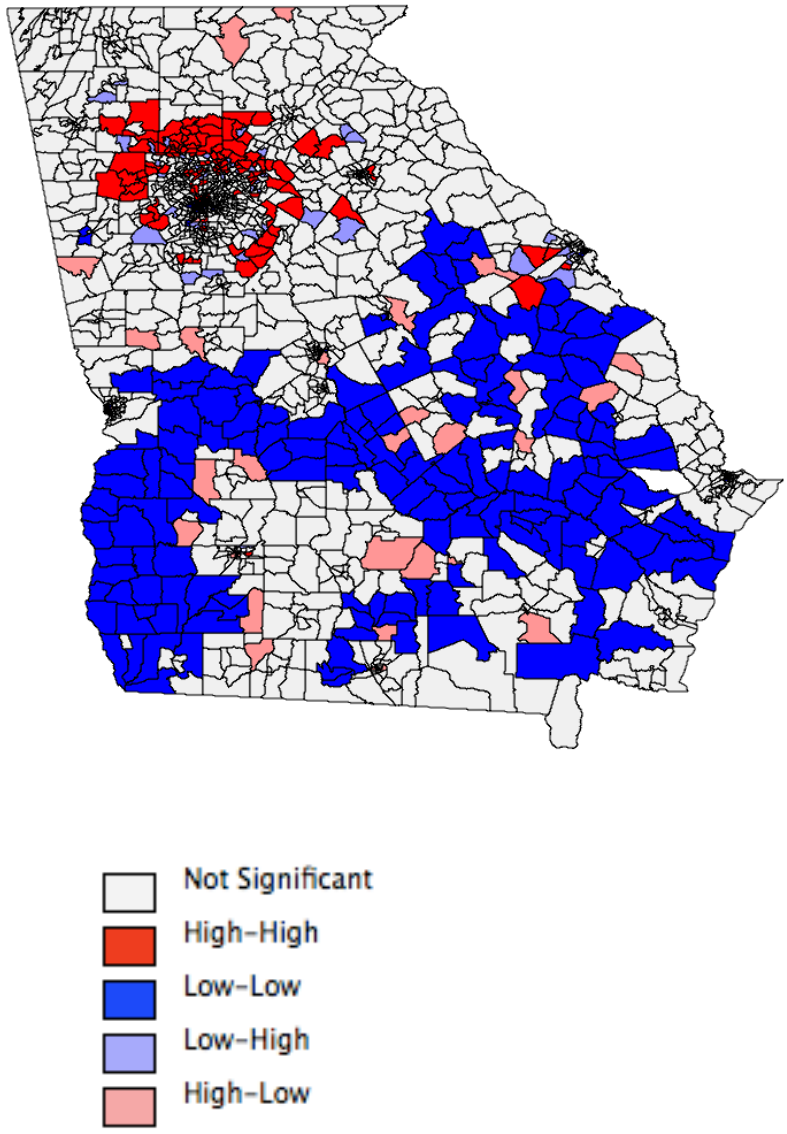


Figure 29. Cluster map of NHL-B SEB-smoothed SIRs

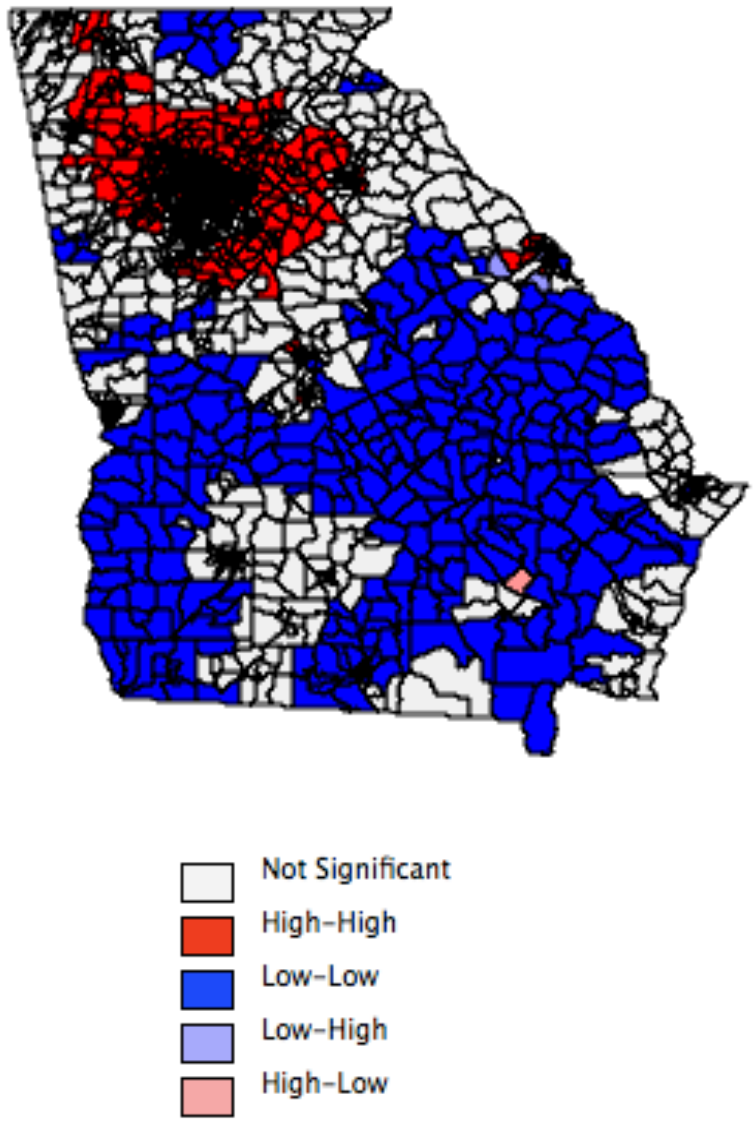


Figure 30. Cluster map of DLBCL SIRs

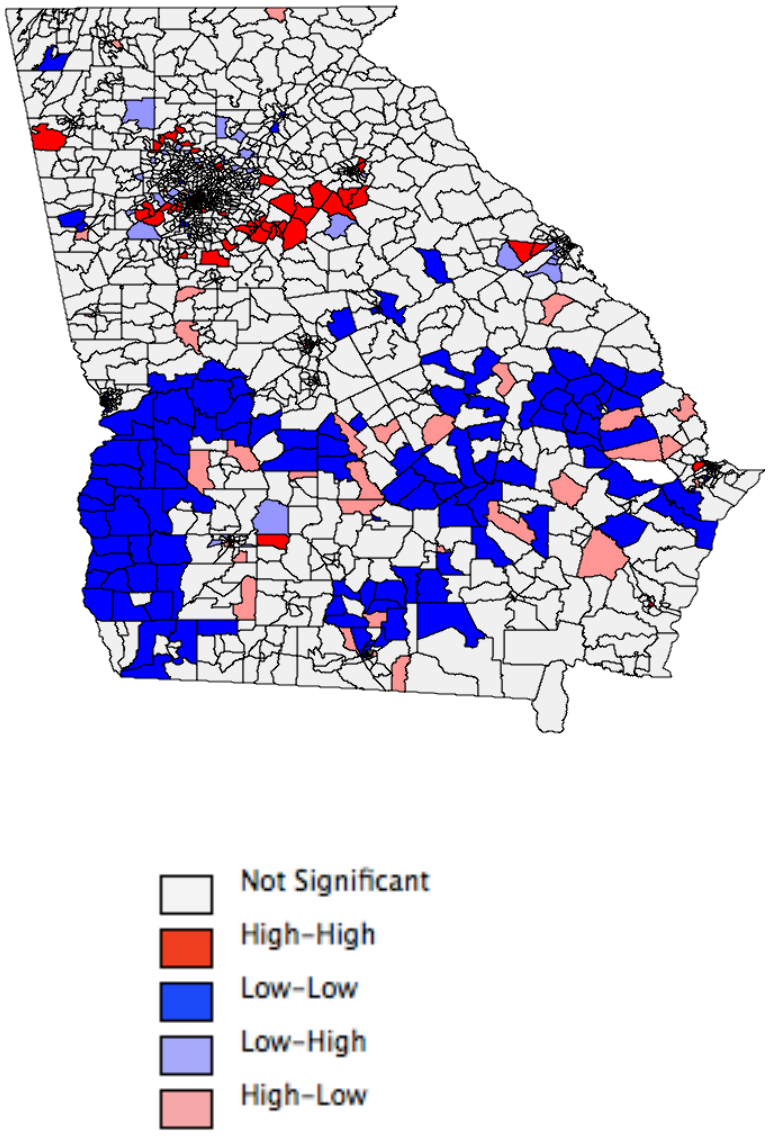


Figure 31. Cluster map of DLBCL SEB-smoothed SIRs

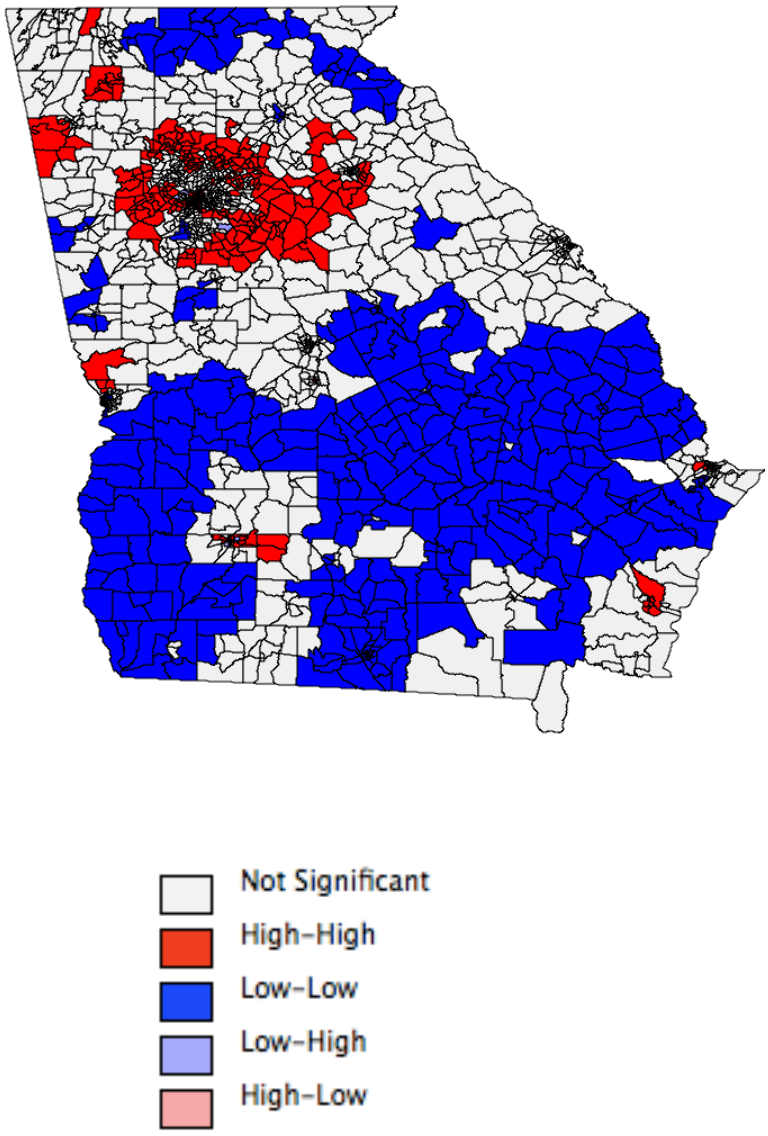


Figure 32. Cluster map of FL SIRs

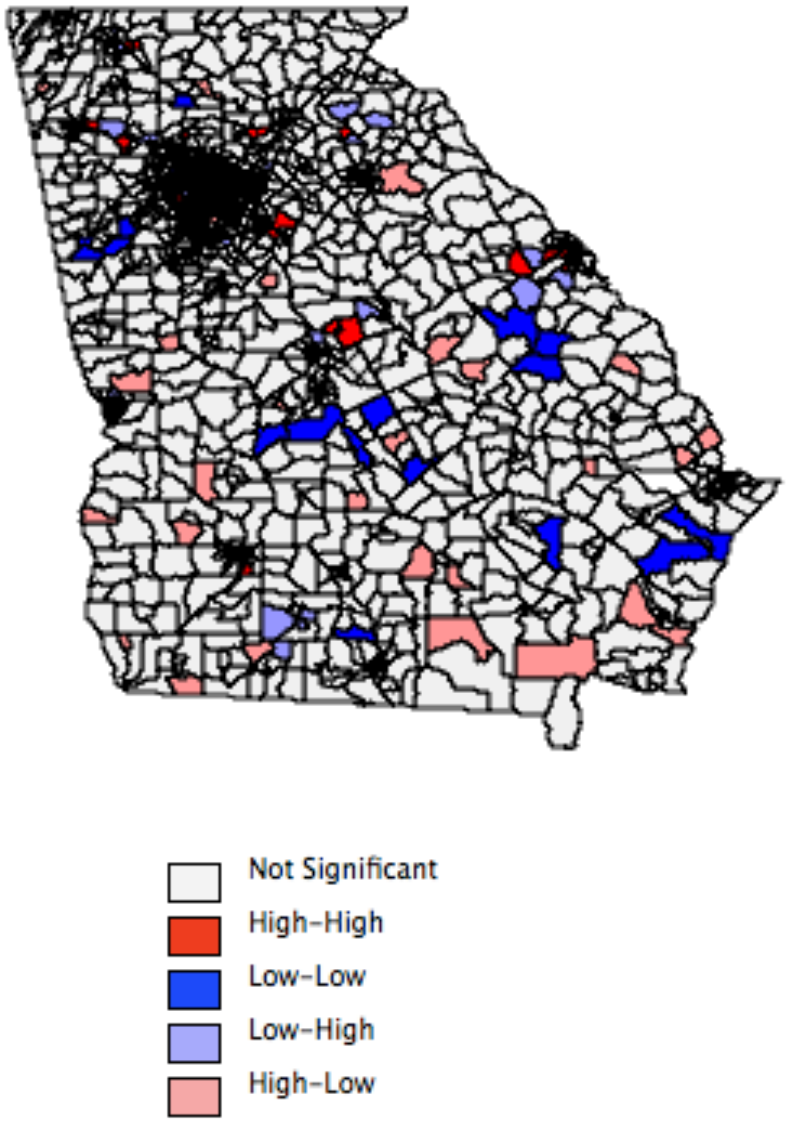


Figure 33. Cluster map of FL SEB-smoothed SIRs

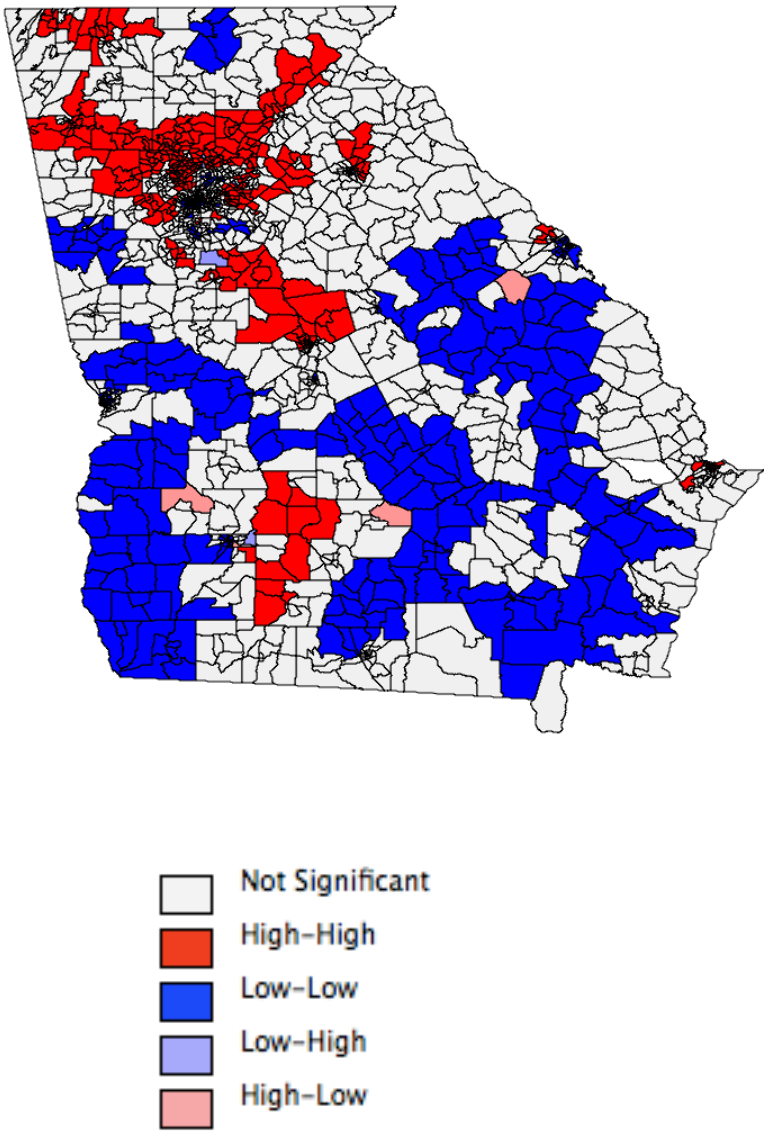


Figure 34. Cluster map of NHL-T SIRs

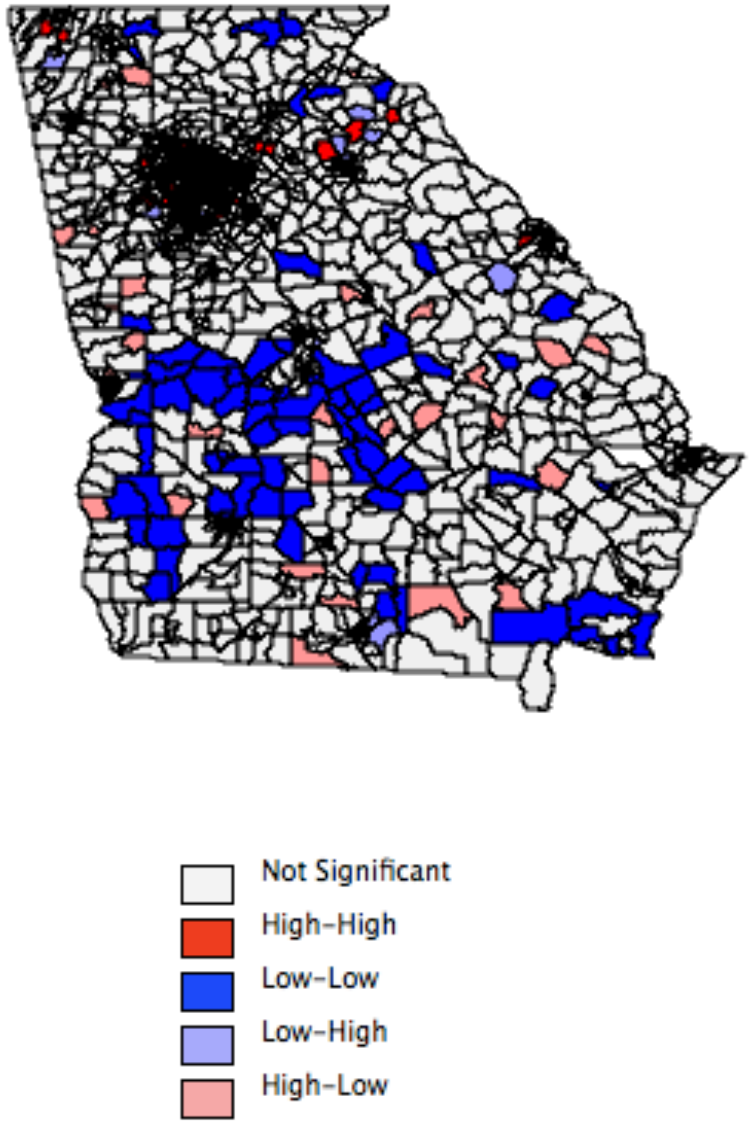


Figure 35. Cluster map of NHL-T SEB-smoothed SIRs

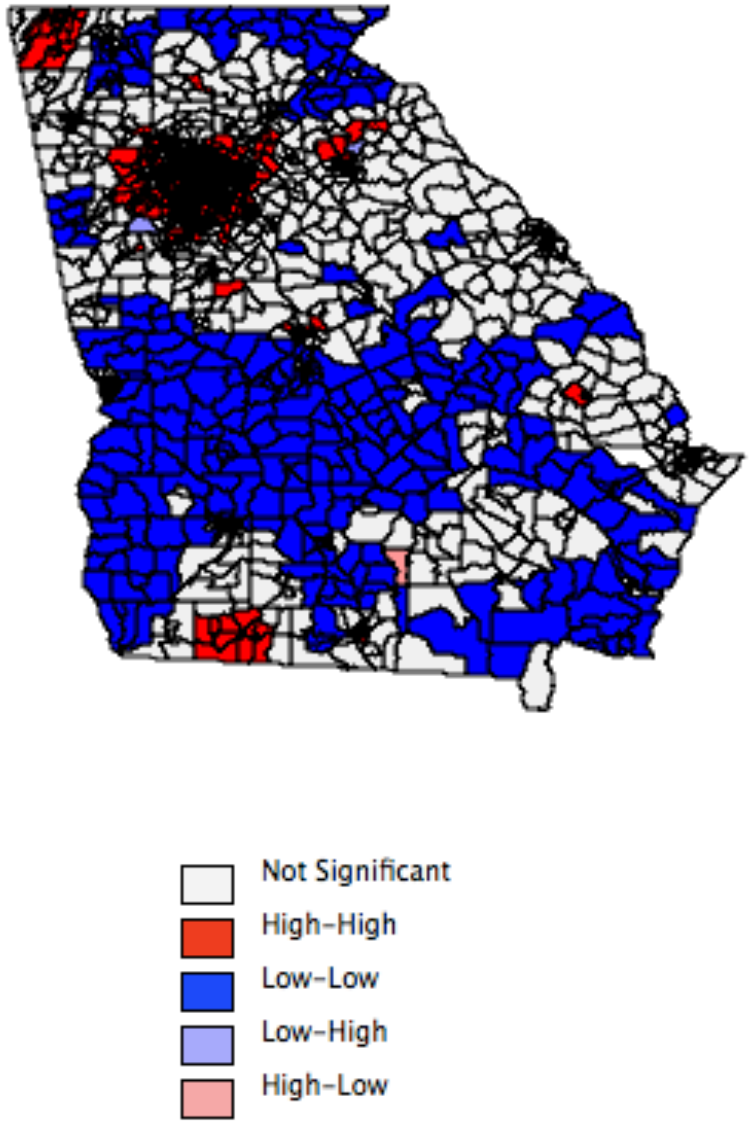


Figure 36. Cluster map of CLL/SLL SIRs

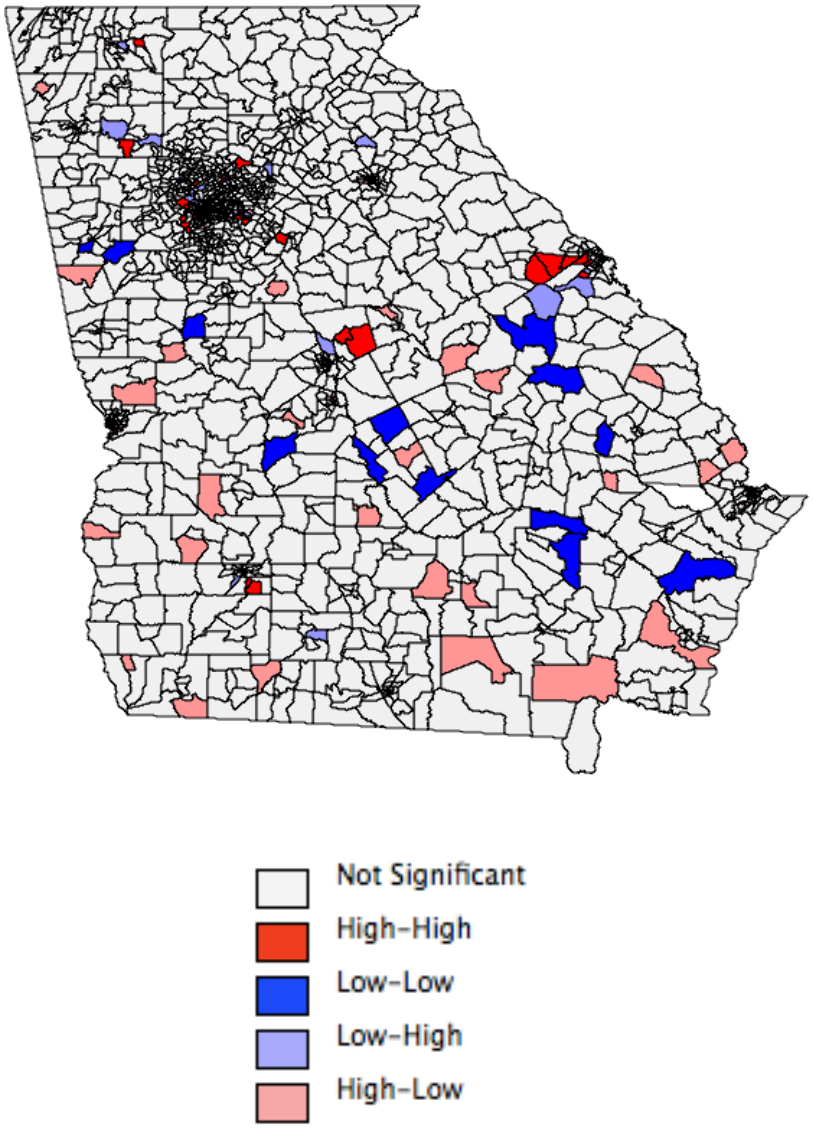


Figure 37. Cluster map of CLL/SLL SEB-smoothed SIRs

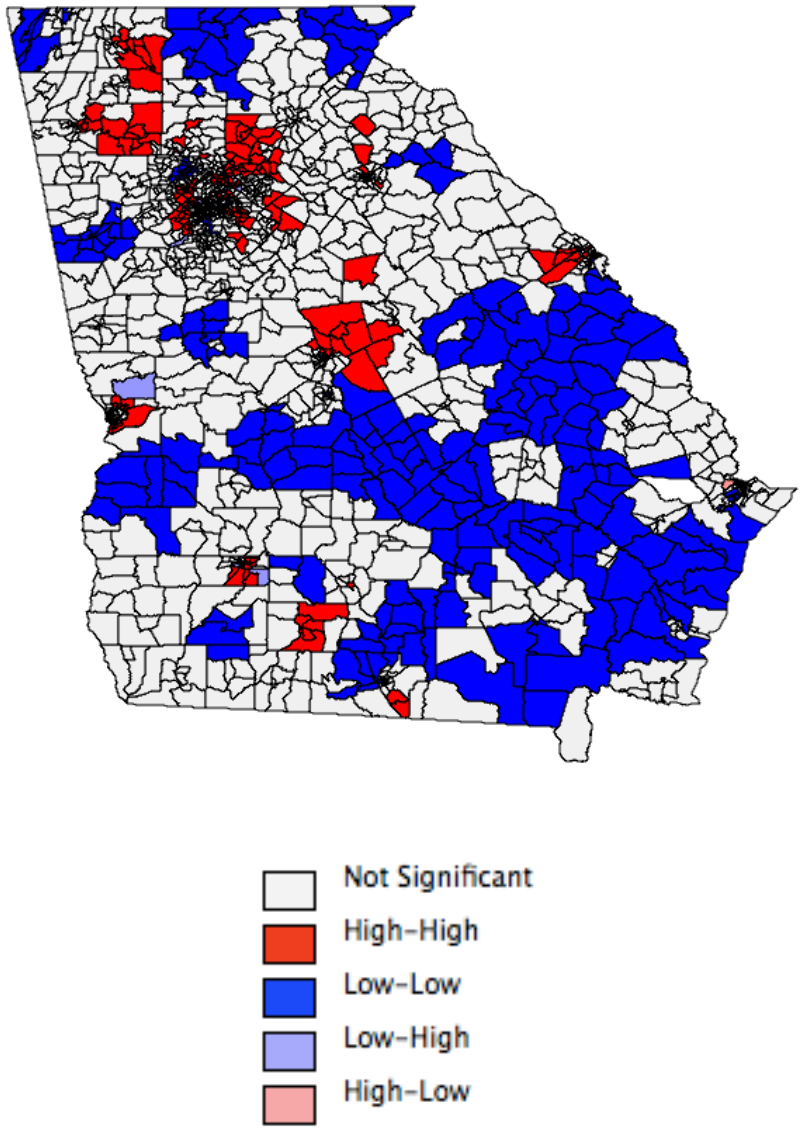
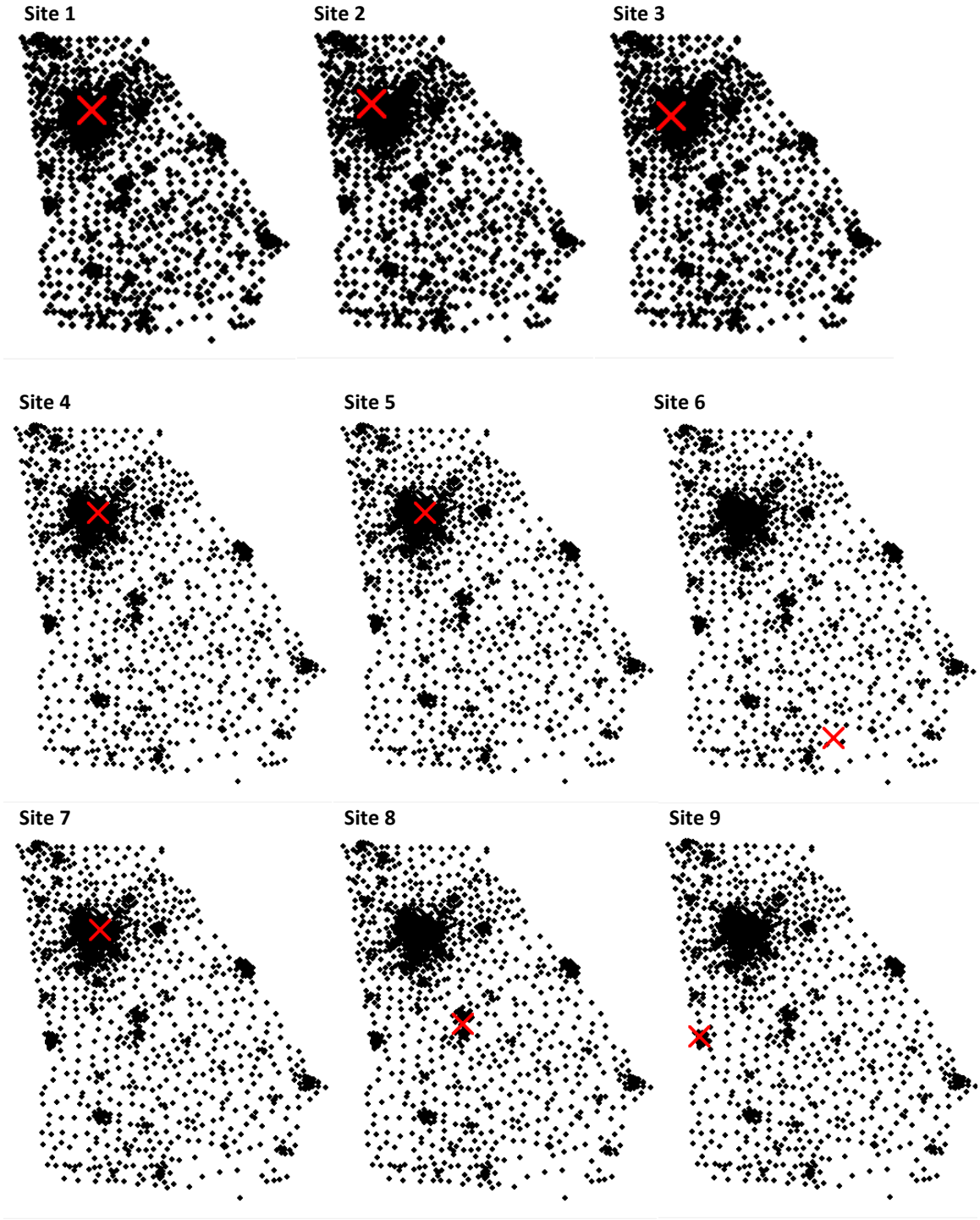
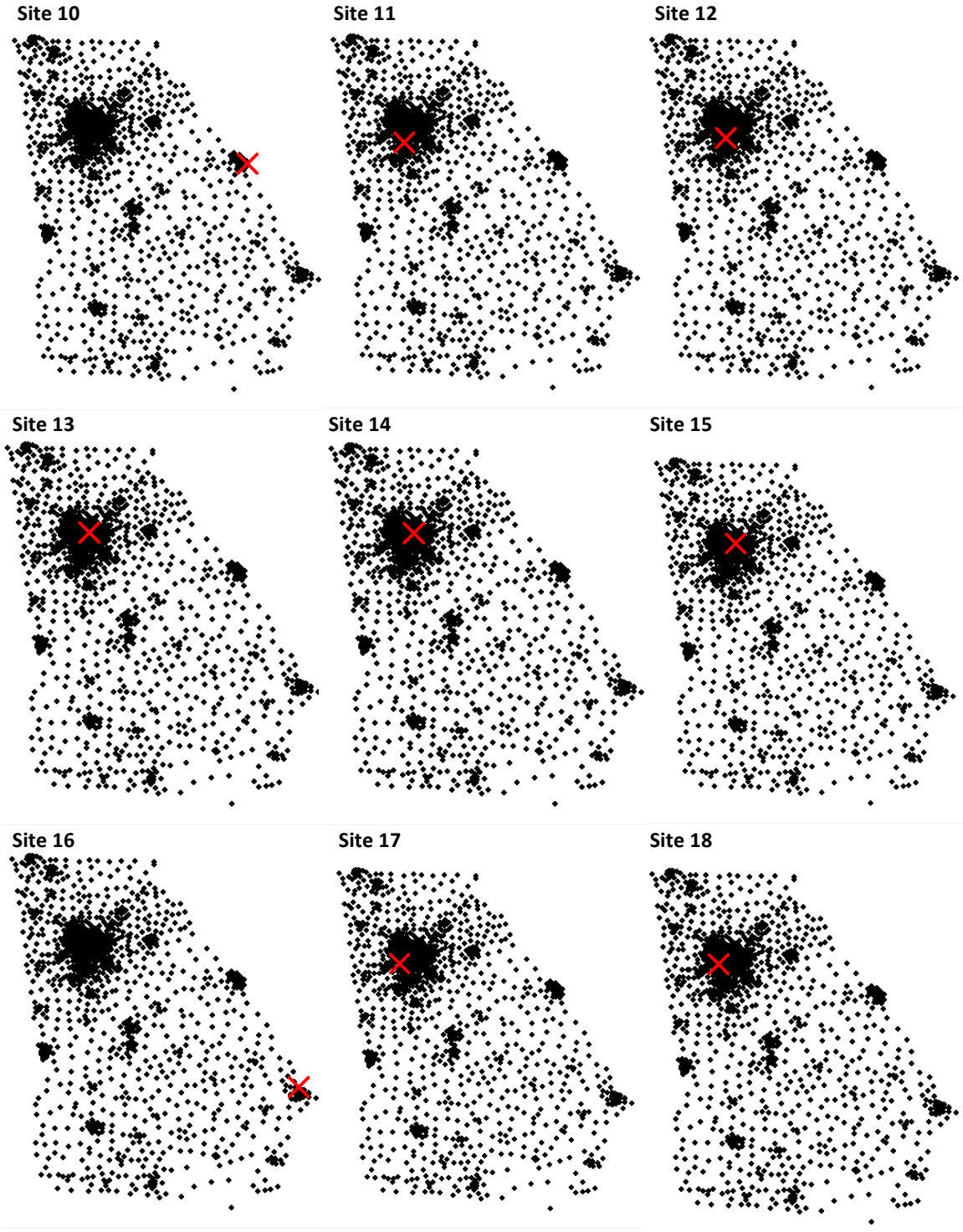


Figure 38. Benzene release sites in Georgia, 1988-1998




Site 10
Site 11
Site 12
Site 13
Site 14
Site 15
Site 16
Site 17
Site 18

Site 19