Geographic Analysis of Childhood Lead Exposure in New York State

Thomas O Talbot, MSPH,* Steven P Forand, MA, Valerie B Haley, MS
Geographic Research and Analysis Section, Bureau of Environmental and Occupational Epidemiology, New York State Department of Health, Troy, NY

Abstract

This study examines the geographic variation in the blood lead levels (BLLs) of New York State children using spatial filtering, contour mapping, and regression techniques. Data for 364,917 children tested for BLLs prior to age two were extracted from New York State's electronic blood lead reporting system. Spatial filtering methods were used to determine which areas of the state had the highest prevalence of children with elevated BLLs (BLLs $\geq$10 µg/dL). The method used a variable filter size to allow for the simultaneous evaluation of urban and rural areas of the state. The results showed that several upstate urban areas had the highest proportion of children with elevated BLLs. Screening rates were also found to be higher in areas with a high proportion of children with elevated BLLs, indicating that areas with a high risk of lead exposure were well screened. Multiple regression analysis, using areas made up of merged zip code regions as the units of observation, was conducted to describe the relationship between the prevalence of children with elevated BLLs and community characteristics. High prevalence of elevated BLLs was predicted in areas with older housing stock, a smaller proportion of high school graduates, and a larger proportion of black births. Separate models were developed for New York City and the rest of the state, since the effect of the variables was lower in New York City.

Keywords: lead poisoning, disease surveillance, socioeconomic status, New York State, children

Introduction

Lead poisoning is considered to be one of the most prevalent and preventable childhood health problems in New York State (1). Blood lead levels (BLLs) as low as 10 micrograms per deciliter (µg/dL) are associated with adverse effects on learning, behavior, and growth. Higher BLLs can lead to anemia, severe central nervous system damage, and even death (2). Young children are at a heightened risk for elevated BLLs because lead intake as a proportion of body mass and metabolic uptake rates are higher in children than in adults. In addition, the central nervous systems of children are more vulnerable during early childhood development (3). Normal mouthing activity may also result in the ingestion of contaminated dust and soil.

A major source of lead exposure for children is lead-based paint in older houses. Indoor house dust may be lead-contaminated when the paint is chipped, peeling, deteriorating, or spread during renovation. Children in New York State are at particular risk because the state has the largest proportion (47%) and largest number of housing units built before 1950 (3.4 million units) of any state. Other sources of lead exposure

* Thomas O Talbot, New York State Department of Health, 547 River Street, Room 200, Troy, NY 12180-2216 USA; (p) 518-402-7950; (f) 518-458-7959; E-mail: tot01@health.state.ny.us

include lead in soil and dust from external paint, industrial emissions, and gasoline; lead in water pipes with lead solder; and lead brought into the home from occupational exposures, hobbies, and ceramic ware.

The prevalence of children with elevated BLLs is not evenly distributed across New York State. Locating areas with a high prevalence of children with elevated BLLs is important for identifying possible exposures in those areas. Two methods were used to examine the geographic variation in the prevalence of elevated BLLs across New York. One method used spatial filtering techniques to identify and display these areas. The other method involved using multiple regression techniques to identify community characteristics associated with areas with high prevalence of elevated lead values.

Displaying rates of disease over a large geographic area has always been problematic. State health departments traditionally display health outcome rates at the county level. These maps may not be very informative; it is difficult to identify high-incidence areas, which may either be localized within part of a county or cross county boundaries. Mapping health outcome data at smaller geographic levels, such as census tract or zip code, often produces rates of disease that vary widely due to chance alone, especially when the health outcome under analysis is rare. In addition, data regarding the underlying population size are often obscured when displayed in this manner.

Population density varies widely across the state. As expected, the distribution of health outcomes follows the population distribution. It therefore becomes necessary to take into account varying population densities when displaying the data on a statewide basis. To accomplish this we used spatial filtering techniques that control the size of the population for which disease rates are estimated. Stabilized rates were then obtained and displayed as a continuous distribution across the state. The geographic patterns in BLLs and blood lead screening were mapped. These maps are useful for identifying areas with high prevalence of exposed children and areas where blood lead screening rates are low.

We also developed a regression model to predict prevalence of elevated BLLs in areas for which insufficient screening data exist. Regression analyses can also identify areas with higher rates of elevated BLLs than we would expect from our model. These areas can be examined more carefully to better characterize the factors that contribute to lead exposure in communities.

Elevated childhood BLLs have been associated with housing and sociodemographic characteristics including older housing stock, a higher proportion of children living below the poverty level, and a lower proportion of high school graduates (4–8). They have also been associated with a higher proportion of households headed by a female, a higher percentage of minority births, and higher population density. Studies have also shown that children's lead levels tend to be higher in the summer months (9–11). Multiple regression techniques were used to examine the relationship between housing and sociodemographic characteristics and children's BLLs across the state. The results from the regression analysis were also examined geographically and compared with the results from the spatial filtering methods.

The objectives of these analyses were to:

- Identify geographic areas of New York State with high prevalence of elevated BLLs in children.
- Identify areas with low blood lead screening rates in children.

- Develop a model describing the prevalence of high BLLs as a function of community characteristics using multiple regression analysis.
- Identify communities that could be targeted with additional educational, screening, and remediation programs.

Materials and Methods

Data

New York State requires universal screening for lead in children under the age of six. The New York State Department of Health currently receives approximately 870,000 blood lead reports a year for both children and adults. Blood lead reports include the name and address of the child, the name and telephone number of a parent or guardian, the blood lead value, and the method by which the blood sample was obtained. A total of 537,704 records for children who were born in 1994 and 1995, resided in New York State, and were screened for blood lead at least once prior to age two were extracted from the lead reporting system (12). This cohort represents 69.3% of the births in New York State in 1994 and 1995.

For children who were screened more than once prior to age two, the highest BLL measure taken by venipuncture was used because these samples are less susceptible to environmental contamination than finger stick samples (13). Finger stick measures were used if no venipuncture measures were available for the child. BLLs were categorized into two groups: less than 10 µg/dL, and 10 µg/dL and above. This is the intervention level recommended by the federal Centers for Disease Control and Prevention. The data were then aggregated at the zip code level. In cases where the child had a missing or invalid zip code, address-matching software (14) was used to assign the correct zip code. To obtain valid zip codes in cases where no street or town information was available for the child, the parent or guardian's phone number was matched to digital phone directories (15). Valid zip codes could not be obtained for 3.8% of the children screened. The remaining dataset contained 364,917 children with blood lead tests, which represented 66.7% of children under two years old born in 1994 and 1995.

To obtain a denominator for screening rates, data on all births in New York State in 1994 and 1995 were obtained from the New York State Bureau of Vital Statistics (16). The data were aggregated at the zip code level and the percent of children tested and percent of black births were calculated. The dataset was then merged with housing and demographic data from the 1990 Census (17) for use in the regression analysis.

Spatial Filtering Methods

In this analysis, it was not practical to map the lead prevalence of the birth cohorts at the zip code level because many of the zip codes had few blood lead tests. Of the 1,601 zip codes, almost half had fewer than 50 children tested for blood lead. Spatial filtering techniques were developed to overcome some of the difficulties in mapping the data by small geographic area. These methods are a variation on the techniques described by Openshaw et al. (18,19), Turnbull et al. (20), and Rushton and Lolonis (21). In the first step, a layer of grid points 1 kilometer apart was created covering the entire state. The zip code file containing the number of births, the total number of lead tests, and the number of lead tests with results higher than 10 µg/dL was mapped, and the zip code centroids were overlaid on the grid point file. Because there are large fluctuations in population density between the urban and rural portions of the state, a stable population denominator was chosen on which to base the screening rates. In this case, rates were based on a minimum of 200 children screened.

The following procedure was then carried out for each grid point. First, the nearest zip code centroid to the grid point was located. The number of children screened for lead and the number of children with elevated lead values were then tabulated. If the number of children screened was less than 200, the next nearest zip code centroid was located and the number of children screened and the number of children with high lead values were added to the values from the previous zip code. This process was continued until the total number of children screened reached 200, at which point the percent of children with elevated lead values was calculated for that grid point. This process was then carried out on each grid point until all grid points in the state had been assigned prevalence rates. Because the grid points were spaced more closely together than the zip code centroids, there was a significant overlap in the area sampled around one grid point and that of its neighboring grid point. To ensure that all information was used, a minimum radius of 0.75 times the spacing of the grid (in this case, 0.75×1.0 km=0.75 km), was used at each grid point. This radius was then extended, when required, to ensure that the minimum number of observations was captured at each grid point.

Contour modeling software (22) in conjunction with desktop mapping software (23) was used to create a contour model based on the percentage of children with elevated lead values at each grid point calculated in the previous step. In this way, it was possible to represent the data continuously as a moving average.

The percentage of children screened for lead was analyzed using a similar methodology. The New York State vital statistics file provided the number of children born in 1994 and 1995 in each zip code. The percentage of children screened at each grid point was calculated based on a minimum of 200 live births.

Regression Analysis Methods

Least squares regression was used to examine the relationship between children's BLLs and community characteristics. Housing variables examined were percent of houses built before 1940, percent of houses built before 1950, and percent of houses vacant. Socioeconomic variables examined were percent of adults age 25 and older who graduated from high school, percent of children under 5 years living below the poverty level, percent Hispanic, percent black births, percent of population that rents a home, and population density. An additional variable, the percent of children in each zip code who were tested in summer and fall (June to November), was also examined.

The percentage of children with elevated BLLs (i.e., $\geq$10 µg/dL) in each zip code, rather than the mean or geometric mean of the BLLs, was chosen as the dependent variable. This was necessary because labs report different detection limits of BLL and extreme results may be due to child-specific traits, such as pica, that could not be controlled for in this type of analysis. The dependent variable was log-transformed to normalize the distribution.

Zip code areas were used as the level of analysis. The distribution of the prevalence of high BLLs in the 1,601 zip codes was found to be bimodal due to a large number of zip codes having no elevated BLLs. This occurred primarily in areas of low population.

To adjust for this, zip codes with demographically similar adjacent neighbors were manually merged to create zip code groups with a minimum number of children screened. All zip codes with fewer than 100 children tested were merged; merging beyond this point provided greatly diminished returns in correlation. The estimates of the rates of elevated BLLs improved and the zero values changed to small values as the data were aggregated. The distribution of the BLLs also changed from a bimodal distribution to a normal distribution. The final number of zip code groups in the dataset used in the regression analysis was 740.

The regression models were developed using SAS software (SAS Institute, Cary, NC) (24). First, the shape and effect of the bivariate associations of each variable with the dependent variable for New York State were examined. Because the effects of the explanatory variables in New York City were muted compared with the effects in the rest of the state, and no available variables explained this difference, two separate models were developed: one for New York City and one (called the Upstate model) for the rest of the state. The maximum R^2 improvement technique was used to find the "best" one-variable model, two-variable model, etc. A parsimonious model was chosen in which the addition of variables would not add significantly more information to the model. The importance of interactions and curvilinearity were then investigated. Diagnostic methods were used to detect influential observations and multicollinearity, verify the linearity of the regression function, and verify the constant variance and normality of the error terms.

The regression residuals were mapped to more rigorously compare the observed and expected rates. The residuals were standardized for the New York City and Upstate models separately so that the variation would be on the same scale for the combined map. The Moran's I statistic was calculated to test for spatial autocorrelation between the regression residuals of neighboring zip codes.

Results

Spatial Filtering Results

The geographic distribution of the prevalence of elevated BLLs is mapped in Figure 1. The map shows that several urban areas of upstate New York have communities with a large portion of children with elevated lead levels; the cities of Buffalo, Rochester, Syracuse, Schenectady, and Albany all had areas in which more than 25% of the children tested had elevated BLLs. The city of Newburgh in Orange County also had a large area in which 20–25% percent of the children tested had elevated BLLs. The high-prevalence areas are small compared with the size of the counties that contain them, but the population densities of these areas are higher than those of their surrounding counties. In no area of New York City did more than 20% of the children have elevated BLLs. There was one small area of Brooklyn in which 15–20% of the children screened had elevated BLLs. The rest of the city had prevalence rates of elevated BLLs that were low compared with upstate urban areas.

The percent of children screened for lead is mapped in Figure 2. In New York State, the areas with the highest prevalence of elevated BLLs ($\geq$25% of children screened with BLLs $\geq$10 μg/dL) also had the highest screening rates, with more than 80% tested. With the exception of the city of Schenectady, all of the major upstate cities with areas hav-

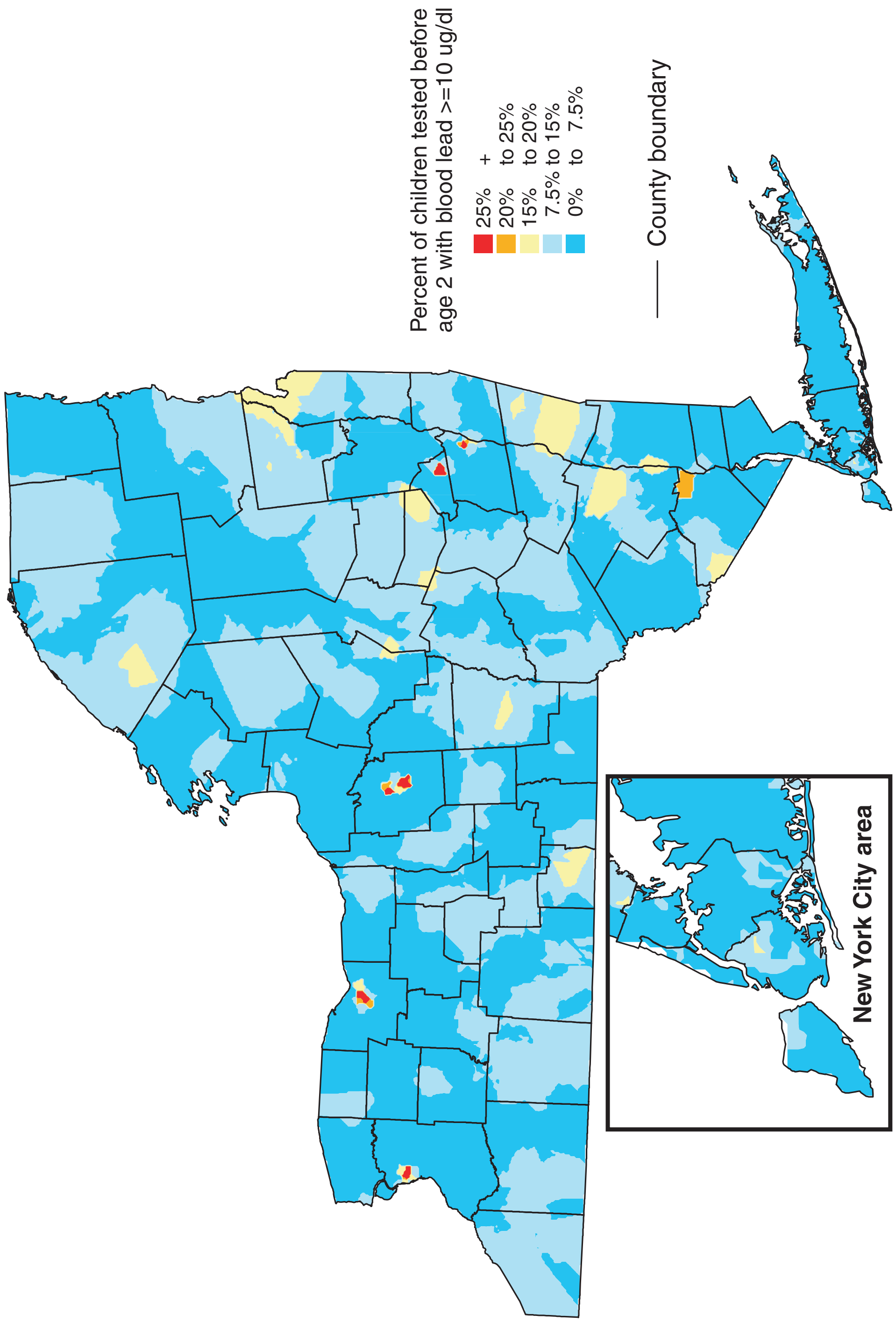

Figure 1 Geographic distribution of blood lead levels in New York State children based on spatial filter method. Rates are based on grid points capturing a minimum of 200 children screened.

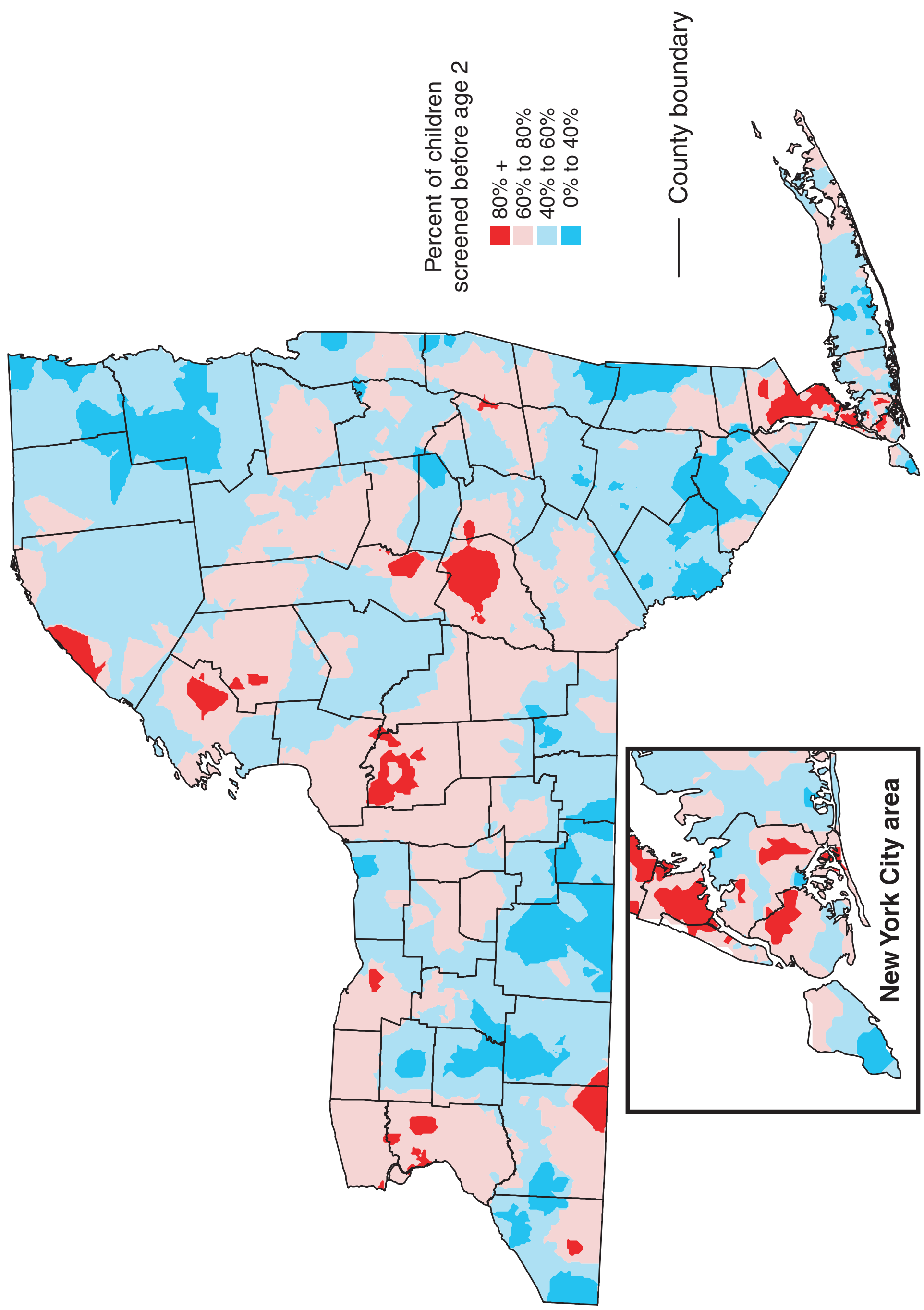

Figure 2 Percent of children screened for blood lead using spatial filter method. Rates based on grid points capturing a minimum of 200 live births.

ing a high prevalence of elevated BLLs also had high screening rates. This trend was also apparent in New York City. More than 80% of the children were screened in the area of Brooklyn that had the highest prevalence of children with elevated BLLs. In addition, high screening rates were also observed in areas of the South Bronx and upper Manhattan.

Regression Analysis Results

Age of housing, race, and education level were the most significant variables in explaining variation in BLLs (Table 1). The variable, percent of housing built before 1940, was selected for inclusion in the model because it was a slightly better predictor of BLL than percent of housing built before 1950. The poverty and education variables were highly correlated ($R=-0.8$), so including both in the model would have been problematic. Education was chosen for inclusion in the final model because it explained more variation. Other demographic variables only explained a small amount of the variation in BLLs, and were not included in the final model.

The model assumptions of linearity of the regression function and constant variance and normality of the error terms were valid. There were no influential observations, and multicollinearity was small.

The observed prevalence of elevated BLLs in the 740 zip code groups is mapped in Figure 3, and the prevalence predicted by the model is shown in Figure 4. The maps show similar patterns.

The regression residuals for the 740 zip code groups are mapped in Figure 5. Areas where more children have elevated BLLs than the model predicts have positive residuals. These areas appear to be clustered in Brooklyn, eastern upstate New York, and eastern Long Island. The Moran's I test statistics show that zip code groups with common boundaries were positively correlated after adjusting for the education, age of housing, and race variables used in the regression model ($p<0.001$). Spatial autocorrelation was also found to be inversely proportional to the distance between zip code group centroids up to a distance of 40 miles ($p<0.001$).

Figure 6 shows the association between the percent of children screened and the percent of children screened with elevated BLLs in the 740 zip code groups. The percent of children screened increases with the percent of children with elevated BLLs. The results were similar to those observed with the spatial filtering method (see Figures 1 and 2). In the areas of the state, excluding New York City, with the highest prevalence of elevated BLLs, 89% of the children had been screened. This effect was also seen in New York City.

Because the independent variable was log-transformed, the meanings of the regression coefficients are more difficult to interpret. Figure 7 contains conditional effect plots to facilitate interpretation of the results. These plots show the predicted value of elevated lead levels versus each of variables used in the model while holding the other variables at their means. The effect of the three major variables (age of housing, education, and race) is stronger in the Upstate model than in New York City model.

Discussion

Spatial filtering techniques were used to identify areas of the state with the highest prevalence of elevated BLLs in children (Figure 1). In addition, we mapped statewide

Table 1 Bivariate and Multivariate Regression Models for Blood Lead Level Analysis in New York City and Upstate New York

	NYC (n=165)[a]			Upstate (n=575)[b]		
	Bivariate Coefficient (adjusted R^2)	**Multivariate Coefficient (standard error)**	**P value**	**Bivariate Coefficient (adjusted R^2)**	**Multivariate Coefficient (standard error)**	**P value**
Intercept		1.6345 (0.1245)			2.3742 (0.2503)	
% housing units built before 1940	0.0088 (0.20)	0.0104 (0.0010)	<0.0001	0.0304 (0.56)	0.0238 (0.0012)	<0.0001
% high school graduates	–0.0123 (0.19)	–0.0064 (0.0015)	<0.0001	–0.0549 (0.36)	–0.0191 (0.0028)	<0.0001
% black births 1994–1995	0.0068 (0.26)	0.0071 (0.0007)	<0.0001	0.0209 (0.18)	0.0093 (0.0013)	<0.0001
% housing units built before 1950	0.0081 (0.19)			0.0028 (0.52)		
% children under five in poverty	0.0082 (0.16)			0.0385 (0.43)		
% renter-occupied housing units	0.0041 (0.06)			0.0268 (0.30)		
% children screened	0.0141 (0.27)			0.0111 (0.06)		
% vacant housing units	–0.0098 (<0.01)			0.0165 (0.06)		
Population density	1.3E–6 (0.04)			1.9E–5 (0.06)		
% Hispanic	0.0033 (0.03)			0.0062 (<0.01)		
% screened (July–November)	–0.0380 (<0.01)			1.6337 (<0.01)		
Adjusted total R^2		0.60			0.64	

[a] NYC model:
ln(% elevated BLL+1)=1.6345+0.0104×(% pre-1940 houses)+.0071×(% black)-0.0064×(% high school grads)

[b] Upstate model:
ln(% elevated BLL+1)=2.3742+0.0238×(% pre-1940 houses)+0.0093×(% black)-0.0191×(% high school grads)

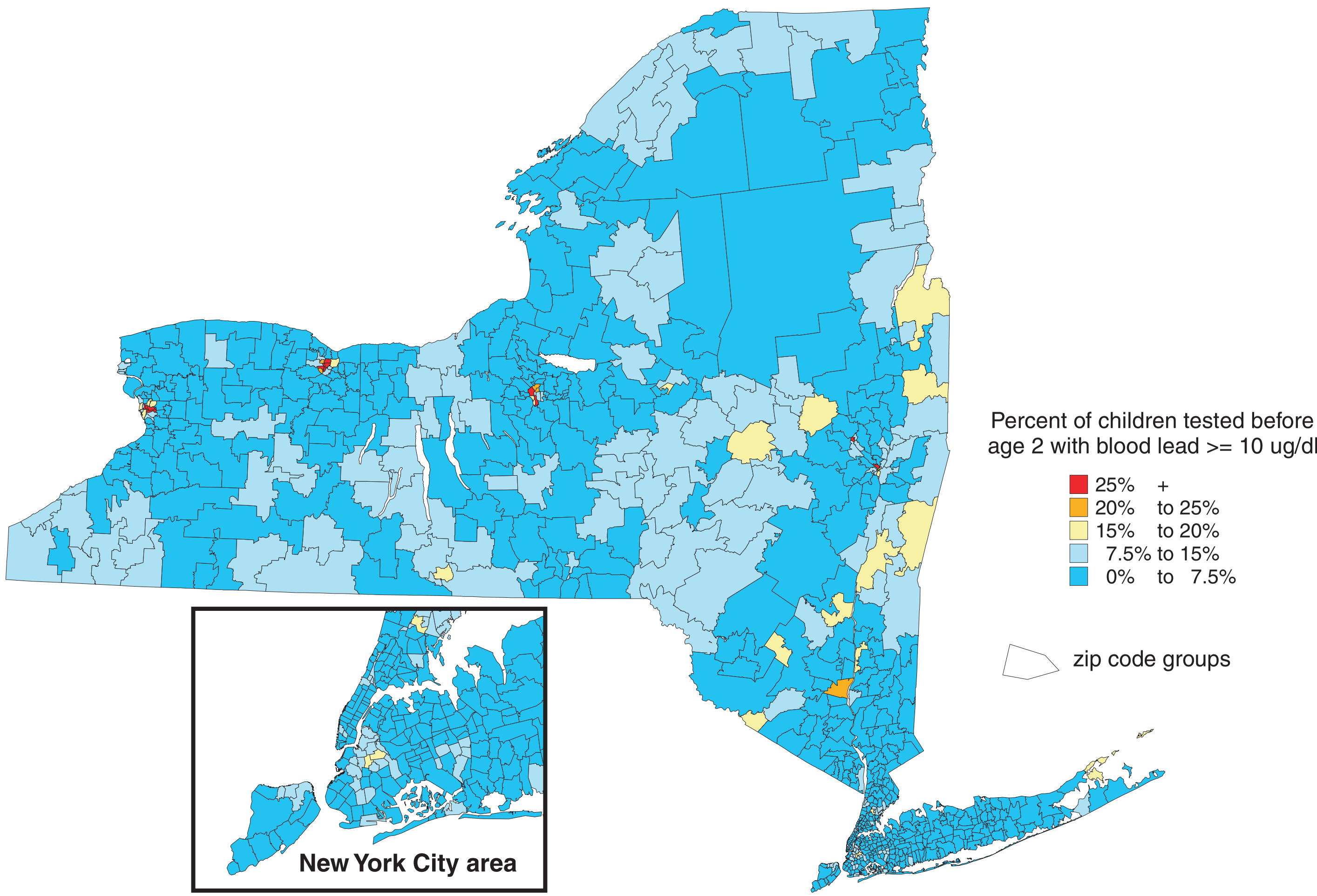

Figure 3 Distribution of elevated blood lead levels in New York State children by zip code group. Areas represent zip code groups with a minimum of 100 children screened.

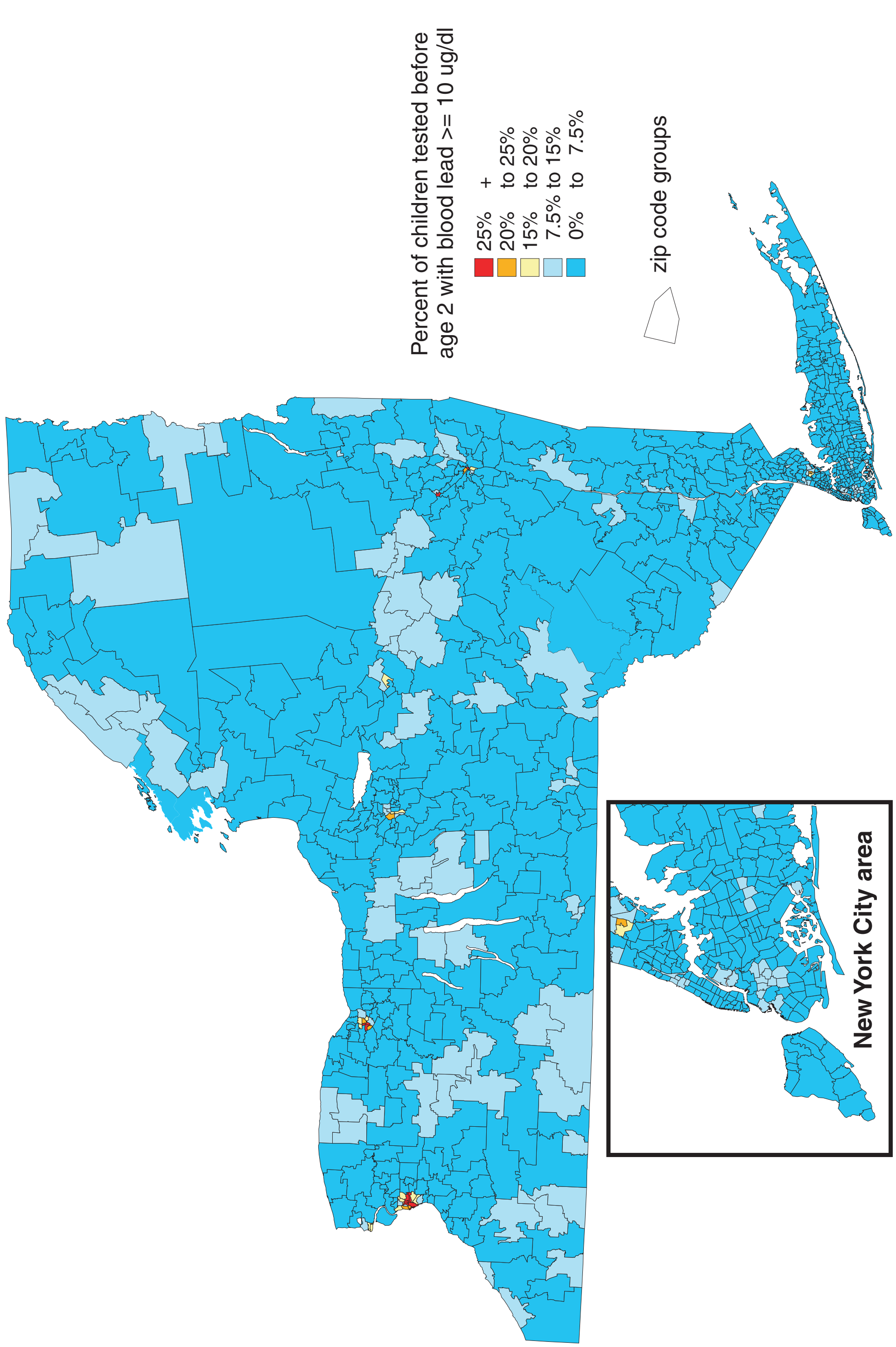

Figure 4 Predicted distribution of children with elevated blood lead levels. Areas represent zip code groups with a minimum of 100 children screened.

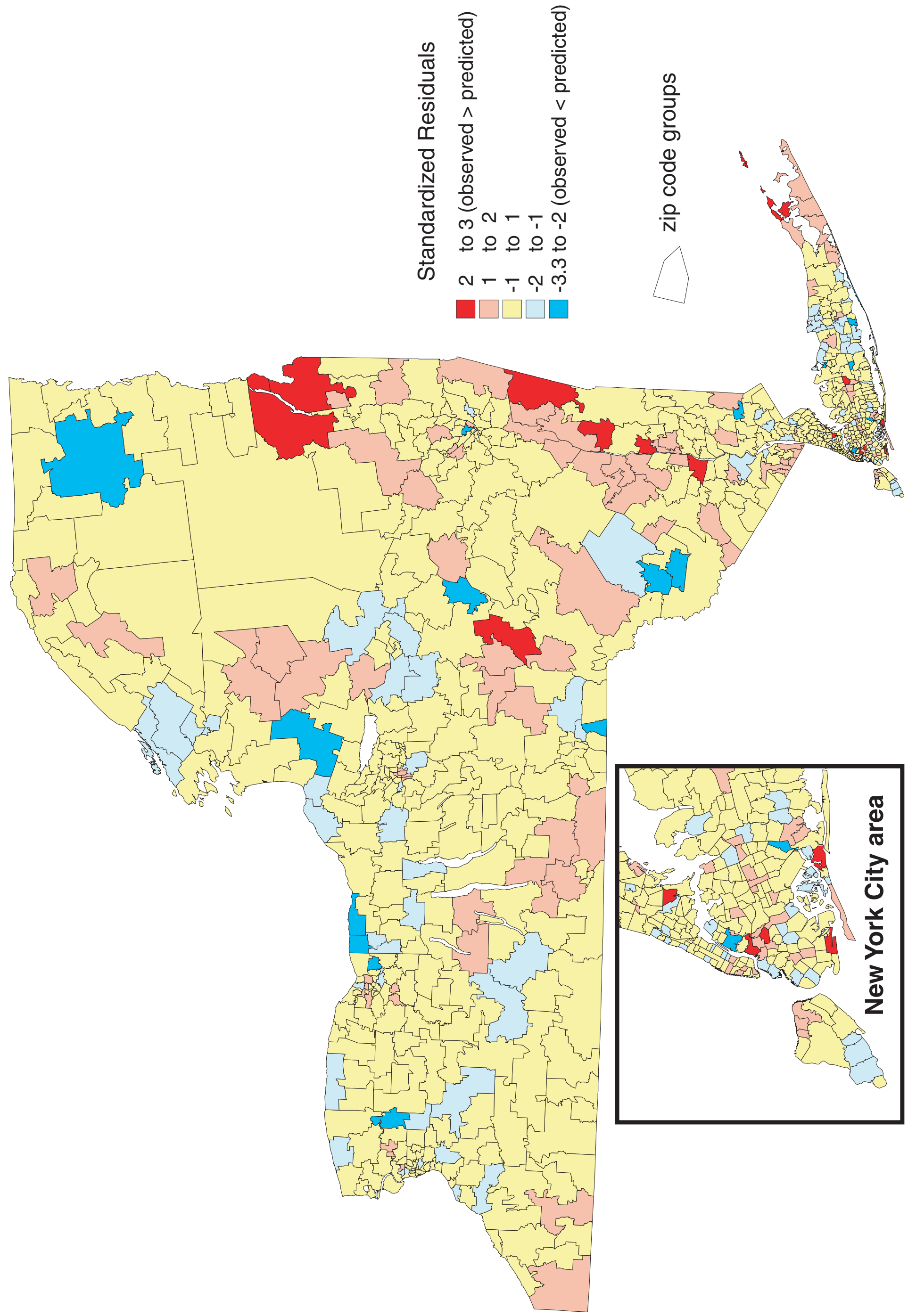

Figure 5 Distribution of regression residuals from the Upstate and New York City models. Areas represent zip code groups with a minimum of 100 children screened.

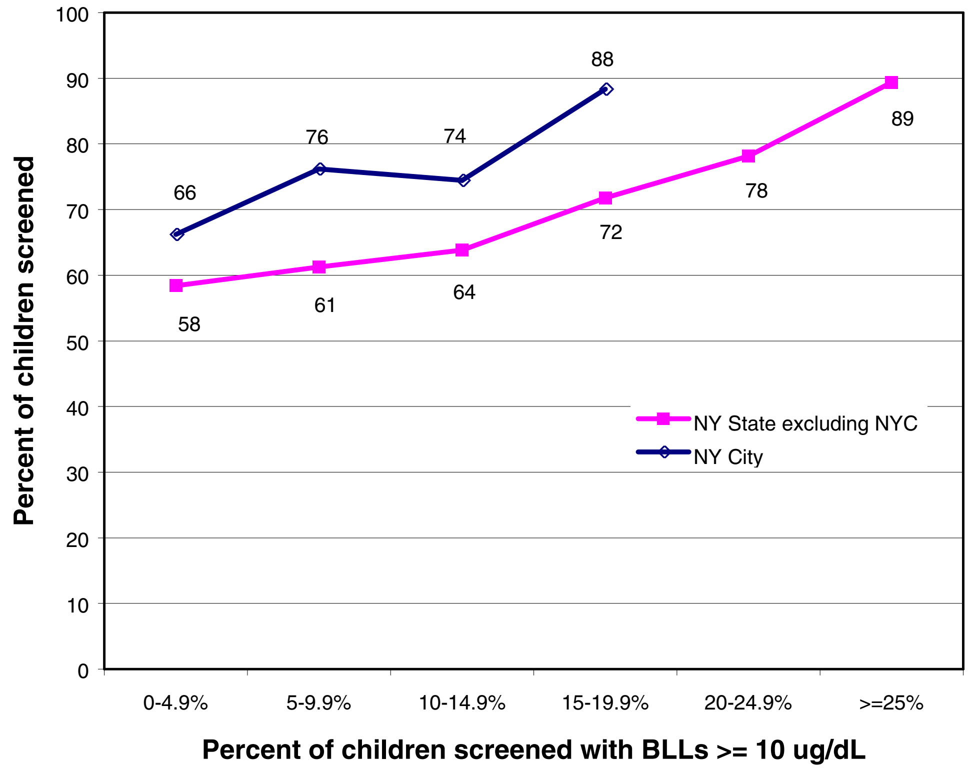

Figure 6 Lead screening and elevated lead rates in children born 1994–1995 and screened prior to age two. Data based on 740 New York State zip code groups. Separate plots are displayed for New York City and the rest of New York State.

screening rates (Figure 2). Screening rates were found to be highest in areas with high proportions of children with elevated BLLs, indicating that areas with a high risk of lead exposure were well screened.

Spatial filtering techniques were used to overcome some of the problems typically seen when mapping disease rates for small areas. Using a spatial filter based on population size rather than a fixed geographic size ensures that enough births were selected at each grid point to calculate a stable rate. One advantage of this method is that we were able to examine the geographic variation in BLLs throughout New York State, which has both rural and urban areas. In areas of high population density the process may capture data from only the nearest zip code. The rate would be stable because the population of that zip code would be large. In areas of low population density it was necessary for each grid point to capture data not only from the nearest zip code, but also from one or more of its neighboring zip codes in order to obtain a stable rate. Because of the varying geographic size of the filter it is possible to display and analyze the data for the entire state using a single summary level.

In addition, we were able to display the information as a continuous distribution

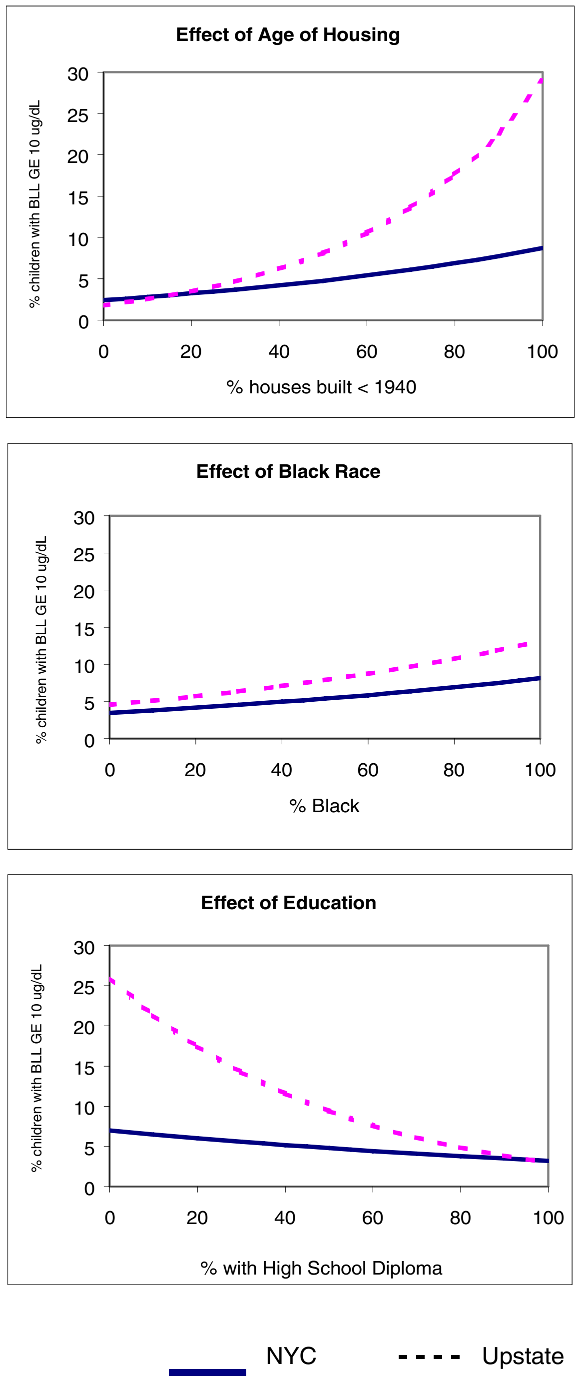

Figure 7 Conditional effect plots for the New York City and Upstate models. Mean values for percent of homes built before 1940=34%, percent black=12%, and percent high school graduates=77%.

rather than one limited to predefined political boundaries. It is illogical to assume that the distribution of a health outcome changes precisely at the border of a county, census tract, or some other predefined boundary. Using grid points spaced much more closely than the unit of analysis ensures that disease rates will be displayed more accurately as a moving average across political boundaries. The effects of the smoothing are critical in areas of low population density.

Multiple regression analysis identified age of housing, education, and race as the community characteristics associated with elevated BLLs. This is consistent with previous studies examining the socioeconomic characteristics associated with elevated BLLs (5–7). This analysis was carried out at the zip code area level. The smaller the units of analysis, the more homogenous the groups will likely be, and the more accurate the estimate of effect. However, the disadvantage of using small groups is that the estimates of disease rates are unstable. As the zip code areas were merged, the resulting correlation coefficient also increased. The increase in the percent-explained variation must be weighed against loss of information from using more heterogeneous areas.

Figure 3 shows the observed prevalence of elevated BLLs in the 740 zip code areas. This map shows a similar pattern of elevated BLLs in the upstate cities to that shown in Figure 1, which was produced using the spatial filtering method. The spatial filtering method is much faster than the manual merging of zip codes. Using it, one can easily plot maps based on the capture of different numbers of children and compare the results. The spatial filtering method can also be used to map individual-level data for which geocoded data are available.

The lower prevalence of elevated BLLs in New York City than in upstate areas is surprising considering that New York City is among the oldest and most densely populated areas of the state. New York City has a higher percentage of older housing stock and minority births and a lower percentage of high school graduates than does the rest of New York State. Based on these characteristics, New York City would be expected to have one of the highest prevalence rates of elevated BLLs in the state. This was not the case, however, and none of the variables examined in this analysis could explain the findings. One possible explanation for this difference is that New York City was one of the first localities to ban lead-based paint, prohibiting residential use beginning in 1960. Lead paint was not banned in the rest of the state until the federally imposed ban took effect in 1978 (26). Further research is needed to determine what factors have contributed to the city's lower prevalence of elevated BLLs.

Spatial autocorrelation of lead prevalence rates suggests that similarity among nearby areas was not completely accounted for by the variables in the regression model. Neighboring zip code groups may be similar because neighborhoods may cross zip code group boundaries, and because children may visit or use services in nearby areas. Environmental characteristics, housing characteristics, secondary occupational exposures, and behavioral factors may also be similar in neighboring zip code groups. Several areas in the state had higher rates of elevated BLLs than we would expect from our model. Further investigation would help us better characterize the risk factors that may be associated with these differences.

This study has several advantages over previously published geographic analyses of lead exposure. These include a larger number of children screened in the study, a higher screening rate, and the use of direct blood lead measurements rather than indicators such as erythrocyte protoporphyrin. In addition, improved mapping techniques, the inclusion of a wide spectrum of population density, and restriction of the data analysis to children under the age of two allowed a more rigorous analysis of the geographic distribution of elevated BLLs in children.

References

1. New York State Lead Poisoning Prevention Advisory Council. 1994. *Annual report.* Albany, NY: New York State Department of Health.
2. Centers for Disease Control and Prevention (CDC). 1997. *Screening young children for lead poisoning: Guidance for state and local public health officials.* Atlanta: CDC.
3. Mushak P. 1992. Defining lead as the premiere environmental health issue for children in America: Criteria and their quantitative application. *Environmental Research* 59:281–309.
4. Bailey A, Sargent JD, Goodman D. 1994. Poisoned landscapes: The epidemiology of environmental lead exposure in Massachusetts children 1990–1991. *Social Science and Medicine* 39(6):757–66.
5. Sargent JD, Bailey A, Simon P, Blake M, Dalton MA. 1997. Census tract analysis of lead exposure in Rhode Island children. *Environmental Research* 74:159–68.
6. Sargent JD, Brown MJ, Freeman JL, Bailey A, Goodman D, Freeman DH. 1995. Childhood lead poisoning in Massachusetts communities: Its association with sociodemographic and housing characteristics. *American Journal of Public Health* 85(4):528–34.
7. Lanphear B, Byrd R. 1998. Community characteristics associated with elevated blood lead levels in children. *Pediatrics* 101(2):264–71.
8. McDade K. 1994. *Factors that relate to children at high risk for elevated blood lead levels in Syracuse, New York.* Thesis. SUNY College of Environmental Science and Forestry, Syracuse, NY.
9. Hunter JM. 1978. The summer disease, some field evidence on seasonality in childhood lead poisoning. *Social Science and Medicine* 12:85–94.
10. Stark AD, Quah RF, Meigs JW, De Louise ER. 1980. Season as a factor in variability of blood-lead levels in children. *Connecticut Medicine* 44(7):410–3.
11. US Environmental Protection Agency (EPA). 1995. *Seasonal rhythms of blood lead levels: Boston 1979–1983.* Washington, DC: EPA. EPA 747-R-94-003.
12. New York State Department of Health, Bureau of Occupational Health. 1994–1997. *Lead reporting system.* Albany, NY: New York State Department of Health.
13. Parsons PJ, Reilly AJ, Esernio-Jenssen D. 1997. Screening children exposed to lead: An assessment of capillary blood lead fingerstick test. *Clinical Chemistry* 43(2):302–11.
14. MapInfo Corporation. 1997, 1998. *MapMarker.* Versions 3.0 and 3.5. Troy, NY: MapInfo Corporation.
15. Digital Directory Assistance. 1996. *Phone disc New York and New England.* October. Bethesda, MD: Digital Directory Assistance.
16. New York State Department of Health, Bureau of Vital Statistics. 1994–1995. *Birth statistical files.* Albany, NY: New York State Department of Health.
17. US Census Bureau. 1993. *1990 decennial census of population and housing, summary tape file 3B (STF3B).* Washington DC: US Government Printing Office. July.
18. Openshaw S, Charlton M, Wymer C, Craft A. 1987. A Mark 1 geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographic Information Systems* 1:335–58.

19. Openshaw S, Charlton M, Craft AW, Birch JM. 1988. Investigation of leukaemia clusters by use of a geographical analysis machine. *The Lancet* 272–3.
20. Turnbull BW, Iwano EJ, Burnett WS, Howe HL, Clark LC. 1990. Monitoring for clusters of disease: Application to leukemia incidence in upstate New York. *American Journal of Epidemiology* 132:S136–43.
21. Rushton G, Lolonis P. 1996. Exploratory spatial analysis of birth defect rates in an urban population. *Statistics in Medicine* 15:717–26.
22. Northwood Geoscience. 1996. *Vertical Mapper.* Version 1.5. Nepean, Ontario, Canada.
23. MapInfo Corporation. 1997. *MapInfo Professional.* Version 4.5. Troy, NY: MapInfo Corporation.
24. SAS Institute, Inc. 1997. *SAS.* Version 6.12. Cary, NC: SAS Institute, Inc.
25. Cliff AD, Ord JK. 1975. *Spatial autocorrelation.* London: Pion.
26. New York State Legislative Commission on Toxic Substances and Hazardous Wastes. 1992. *Hearing report.* March.