Spatial Patterns and Health Disparities in Pediatric Lead Exposure in Chicago: Characteristics and Profiles of High-Risk Neighborhoods

Tonny J. Oyana \textsuperscript{a} & Florence M. Margai \textsuperscript{b}

\textsuperscript{a} Southern Illinois University
\textsuperscript{b} Binghamton University, State University of New York

Published online: 04 Dec 2009.


To link to this article: http://dx.doi.org/10.1080/00330120903375894
Spatial Patterns and Health Disparities in Pediatric Lead Exposure in Chicago: Characteristics and Profiles of High-Risk Neighborhoods

Tonny J. Oyana
Southern Illinois University

Florence M. Margai
Binghamton University, State University of New York

Lead poisoning remains a major environmental health threat and a persistent source of health disparities in the United States. In this retrospective study, statistical and geospatial approaches were used to evaluate age- and gender-specific differences in childhood lead prevalence across Chicago, assess the spatiotemporal dynamics of the disease, and identify the socioeconomic and racial composition of high-risk communities. Elevated blood lead levels (≥10 µg/dL of lead) decreased significantly during the study period but disparities persisted across neighborhoods. A significant association was observed between high-risk neighborhoods and housing age, low income, and minority populations. These findings provide insights into the complex geographies of lead exposure and could serve as a basis for developing more targeted health intervention programs.

Key Words: environmental health, geostatistical analysis, health disparities, lead poisoning, spatial analysis.

El envenenamiento con plomo sigue siendo una de las principales amenazas ambientales para la salud y una fuente persistente de las disparidades en materia de salubridad en los Estados Unidos. En este estudio retrospectivo se usaron enfoques estadísticos y geoespaciales para evaluar las diferencias específicas por edad y género en la prevalencia del plomo en la niñez de Chicago, estimar la dinámica espacio-temporal de la enfermedad, e identificar la composición socioeconómica y racial de las comunidades en riesgo alto. Los altos niveles de plomo en la sangre (≥10 µg/dL de plomo) disminuyeron significativamente durante la época del estudio, pero las disparidades se mantuvieron entre los vecindarios. Se observó una asociación significativa entre los vecindarios de alto riesgo y la antigüedad de la vivienda, los bajos ingresos y la población de minorías.

Estos descubrimientos proveen ideas sobre las complejas geografías de exposición al plomo y podrían servir de base para el desarrollo de más programas específicamente orientados hacia la salud. Palabras clave: salud ambiental, análisis geoestadístico, disparidades de la salud, envenenamiento con plomo, análisis espacial.

*This research was approved by the University Human Investigation Review Board in accordance with the guidelines for the protection of human subjects. The research was supported in part by the Southern Illinois University Carbondale Office of Research, Development and Administration Faculty Seed Grant. We wish to thank the PG Editor, Sharmistha Bagchi-Sen, and the anonymous reviewers for their comments on earlier versions of this article. We appreciate the assistance of David Burke for the initial processing and analysis of the data. Kara Scott assisted with literature review and proofread the first draft. Data pertaining to pediatric lead exposure were obtained from the Chicago Department of Public Health with the assistance of Anne Evens and Patrick MacRoy.
One of the most persistent indicators of racial and ethnic health disparities in the United States is environmental lead exposure and the deleterious health outcomes among residents in low-income and minority-dominated communities. Exposure to lead contamination even in low but cumulative doses can potentially damage the human nervous, hematopoietic, endocrine, renal, and reproductive systems (Centers for Disease Control and Prevention [CDC] 1997). Children are the most vulnerable of all population groups, yet it wasn’t until 1967 that pediatric lead poisoning was officially recognized as a major health threat in the United States. Since then, interventive efforts that include early childhood screening, the lowering of the threshold for elevated blood lead levels (BLLs) to 10 µg/dL, the implementation of community awareness programs, and regulatory bans on the usage of lead-based products have all resulted in a continuous reduction in lead poisoning rates (Mushak et al. 1989; Lin-Fu 1992; CDC 1997; Markowitz 2000; Warren 2000; Jacobs, Clickner, and Zhou 2002).

Unfortunately, not all children in the United States have benefited equally from these interventive programs (Griffith et al. 1998). In many localities nationwide, and among some population groups, the risk of childhood lead poisoning remains unacceptably high. For example, the 1999–2002 National Health and Nutritional Survey (NHANES) data released by the CDC showed that non-Hispanic black children, aged one to five years, continue to have the highest prevalence levels of 3.1 percent, with more than twice the observed rate among non-Hispanic white children (CDC 2005). Such disparities are most evident in the nation’s historic and ethnically diverse cities such as Chicago, where endemic levels of lead poisoning have been reported (Brown et al. 2000; Bernard and McGeehin 2003; Oyana and Margai 2007). Approximately 11.2 percent of Chicago’s children are reported to have elevated BLLs, nearly five times the national rate. Further analysis of BLLs obtained from children residing in just two of the inner-city neighborhoods showed that 98 percent of those with elevated BLLs were black and only 0.5 percent were white (Dignam et al. 2004). Racial and ethnic disparities in pediatric lead toxicity abound despite the nationwide progress in reducing the prevalence of elevated BLLs.

Disparities in lead exposure and the negative health outcomes translate into mounting health care costs for individual families and local communities. To date, lead poisoning is the costliest environmental illness in the United States. Landrigan et al. (2002) estimated the total annual costs of lead poisoning at $43.4 billion, nearly twenty times higher than the cost of pediatric asthma and more than 100 times higher than the costs incurred from childhood cancer. In Chicago, Brown et al. (2000) found that within a five-year period (1993–1997), approximately $7.7 million was incurred in hospital charges for caring for lead-poisoned children. Because most of the afflicted children were from socioeconomically disadvantaged households, Medicaid paid for two thirds of those charges.

The costs of caring for children with lead poisoning also extend beyond hospital stays to include long-term irreversible health problems such as learning and behavioral disorders (Mushak et al. 1989; Needleman and Gatsonis 1990; Lanphear et al. 2000; Margai and Henry 2003), decreased mental ability and intelligence quotient (Schwartz and Levin 1991; Bellinger, Stiles, and Needleman 1992; Schwartz 1992, 1994), hearing impairment (Lanphear 2005), decreased attention span (Sciarillo, Alexander, and Farrell 1992; National Academy of Sciences 1993; Tong et al. 1999), delinquency and criminal behavior (Thacker et al. 1992; Canfield et al. 2003), and other nervous system disorders (Mushak et al. 1989; CDC 1991; Lidsky and Schneider 2006).

In light of these challenges facing especially urbanized communities across the United States, the purpose of this study was to document the temporal dynamics and complex geographies of pediatric lead poisoning in Chicago using a large-scale surveillance database. The specific goals were threefold: (1) to document the temporal trends in age- and gender-specific prevalence levels among children; (2) to spatially predict lead prevalence rates and identify local hot spots; and (3) to characterize the disparate patterns of risk within racially diverse neighborhoods for potential use in targeted intervention programs. This study relied on geostatistical approaches to detect local anomalies of the disease. Spatial epidemiological models were used to quantify the risks of childhood lead poisoning across neighborhoods. Later, the racial and
socioeconomic profiles of high-risk neighborhoods were generated statistically based on the census geography.

**Childhood Lead Poisoning: Disparate Risks and Exposure Sources**

Research on lead contamination consistently shows that the toxicity levels and the adverse health effects are more observable in children than in adults (CDC 1997; Jacobs, Clickner, and Zhou 2002; Margai and Henry 2003). The main sources and pathways of childhood exposure and the resulting health effects are summarized in Figure 1. Dust lead and remnants of lead-based paint in the built-up environment are the principal contributors. Lead is activated from these sources when it deteriorates or is disturbed during building maintenance activities. Although there are other exposure sources such as air emissions from industry and transportation and contaminated toys, soil, food, and drinking water, lead paint remains, by far, the most problematic source of pediatric contamination in the United States, primarily because it is the most accessible source for children (Clark et al. 1985; American Academy of Pediatrics 1998; Lanphear, Byrd, et al. 1998; Brown et al. 2000; Markowitz 2000; Jacobs, Clickner, and Zhou 2002; Bashir 2002; Lanphear 2005).

![Diagram](image-url)

**Figure 1** Known sources and pathways of lead exposure and potential health effects.
In many historic cities, lead contamination is persistent because of the predominance of old housing stock and their industrial heritage. For example, in Massachusetts, Bailey et al. (1994) found that twelve communities with the longest history of industrialization accounted for 40 percent of lead poisoning cases. In Syracuse, New York, Griffith et al. (1998) found two distinct geographies of elevated lead to be consistent with historical settlement patterns.

The risks of pediatric lead exposure are heightened further by sociodemographic factors such as age, race, economic status, and renter-occupied housing. In some studies, non-Hispanic blacks and low-income neighborhoods were independent predictors of elevated BLLs in children aged one to five years (CDC 1997; Lanphear 2005). Other studies found childhood lead toxicity to be concentrated in two groups: impoverished children who live in older, poorly maintained rental properties and affluent children whose families renovate and reside in older housing units (Lanphear, Byrd, et al. 1998). It was further demonstrated by Margai and Henry (2003) and is consistent with clinical studies (Shinn et al. 2000; Binns, Kim, and Campbell 2001) that high-risk areas for learning disabilities were synonymous with significant sources of lead poisoning. Collectively, these studies show that patterns of pediatric lead exposure are the result of risk factors associated with the environmental conditions of the child’s residence and socioeconomic disadvantages. A geospatial analysis of the disease and the associative risk factors is therefore relevant in the spatial characterization of neighborhoods of concern.

The use of geospatial methods to better understand the spatial dynamics, characteristics, and risks of childhood lead poisoning is not new, however. Over the years, a great deal of research has been devoted to the geographic distribution of the disease and more generally to the development of spatial and clustering techniques for characterizing the risk areas. Wartenberg (1992), Guthe et al. (1992), Sargent et al. (1997), Margai et al. (1997), and Miranda, Dolinoy, and Overstreet (2002) each proposed unique models for predicting or modeling high-risk areas for childhood lead poisoning. To the best of our knowledge, however, no empirical study has utilized the geostatistical approaches, specifically kriging, to spatially predict and detect the local anomalies of elevated lead levels among children. Nor has any study sought to evaluate the persistence or reduction of these risks over time and among racially diverse communities. Our study sought to accomplish these objectives using a surveillance database that provided a unique opportunity to adequately identify these spatiotemporal patterns and a basis for developing lead intervention campaigns.

Investigating the Risks of Pediatric Lead Exposure in Chicago

Study Area

Empirical data were garnered from Chicago, in Cook County, Illinois. Chicago is an old, industrialized city, stretching approximately 228 square miles along the southwestern shores of Lake Michigan (Figure 2). The estimated population is 2 million, approximately 44.5 percent of whom are white, 34.2 percent are African Americans, and 4.2 percent are Asian American and Pacific Islanders. Self-reported ethnic classifications indicate that 26 percent of the residents are of Hispanic descent. This city, overall, has the second largest black population and the third largest Latino population in the United States, placing it among the most culturally diverse cities in the nation (U.S. Census Bureau 2000).

Chicago has high levels of residential segregation, the fourth highest nationally, with a dissimilarity index of 80.3 among African Americans and whites, and 61.1 between Latinos and whites (McArdle 2002). In Figure 3, showing the distribution of the major racial and ethnic groups, non-Hispanic whites reside mostly in the north around Lake Michigan, the northwestern areas, and suburban outskirts in the far southern and southwestern edges of the city. Black residents also have distinctive geographies, with large concentrations on the city’s Westside and the south. Latino communities are found at the cross-point of three neighborhoods: the northwestern neighborhoods, the lower Westside adjacent to the predominantly black neighborhood, and the northern portion of the Southside. Asians are not well represented in the southern areas of the city where blacks and Latinos reside. Rather, they are fairly well distributed in
areas adjacent to the white communities in the north and northwestern areas. Small clusters of Asian populations are also found close to the downtown areas of the city.

Housing characteristics of the city also show discernable patterns. About 59 percent of Chicago’s housing units (over 600,000) were constructed before 1950, nearly three times the national average of 23 percent (Chicago Department of Public Health [CDPH] 2006). As noted earlier, such homes serve as residual sources of childhood lead exposure. Their distributional patterns, coupled with the inherent complexity of the racial and ethnic geography, make the city an excellent candidate for our study.

**Childhood Blood Lead Level Data**

The research focused on children five years or under, comprising about 9 percent of the total city population (263,486 persons). The BLL data for these children were obtained from the Lead Poison Testing and Prevention Program of the CDPH. This was a large-scale surveillance database compiled by CDPH over a seven-year period (1997–2003). Regulations imposed previously by the state of Illinois required all children between the ages of six months and seven years, particularly those residing in high-risk areas, to be tested before attending a licensed day care or elementary school. Within Chicago, all children were considered to be at high risk; therefore, those aged four years and younger were required to test regularly at intervals of six, twelve, eighteen, twenty-four, and thirty-six months or at ages nine, fifteen, twenty-four, and thirty-six months (CDPH 2006). Further, to promote the early identification of the disease, additional regulations were imposed by the city requiring blood lead screening for every child at least three times before age three and again at age three. Those with BLLs exceeding 10 µg/dL had to be retested within ninety days. These regulations, coupled with relatively high
compliance rates, resulted in a comprehensive database that was deemed by the authors to be spatially and statistically representative of the at-risk population.

The data set, containing records in excess of 881,385, included all reported blood lead screenings for every child tested from January 1, 1997, through December 31, 2003. Of these, 47 percent of the subjects had been tested multiple times for a variety of reasons, including the need to confirm test results as well as meet the guidelines already noted. Due to a

Figure 3  Spatial distribution of racial and ethnic groups in Chicago. The maps show distinct concentrations of racial and ethnic groups in the city. Whites and Asians live in adjacent communities in the north, and blacks and Hispanics occupy areas in the south and west.
rigorous human subject protocol and approval process and the need to protect the personal identity and physical locations of tested children, the initial geocoding of the data was done by the CDPH. With assistance from a student intern, the data were cross-checked, and the geocoded matches were reviewed. The geocoded results were successfully enhanced and matched to slightly over 90 percent. In a few cases where children had missing physical addresses to be used in assigning geographic coordinates, the ZIP code centroid was utilized. All names and addresses of the subjects were removed and replaced with unique identifiers. The final data set consisted of seven variables, including the unique identifier, census block code, age of child, BLL, sex, and two types of blood test samples, based on venous (the preferred diagnostic method) and capillary draws.

Data Analysis

BLL Data Preprocessing and Management

Following acquisition, three preliminary steps were taken to process the data. These included data queries and deduplication, computation of prevalence rates, and data aggregation into multiple census geographies. The unique identifiers in the database made the data management process easier and more efficient. These identifiers were used to track multiple screenings and identify children who had relocated.

The process began by first splitting the original BLL data set into several smaller data tables based on age, sex, and date of blood test. Each of these tables was then deduplicated to account for the children who underwent multiple screenings. This deduplication process allowed for the calculation of the actual number of children screened over the entire study period as well as the number of children screened per year. The deduplication process was performed using ESRI’s ArcView 3.3 Extension based on Avenue scripting Compiled Table Tool to sort the data and identify different sampling tests, age, test year, and gender. This process reduced the data set from approximately 881,385 records to 469,297 records, with each record representing the test result obtained from a single BLL screening per person or the average test result for those with multiple screenings. Seven data tables were created, one for each year in the study period (1997–2003), and within each year, a series of analyses were conducted. In this article, however, we report only on the results obtained for 1997, 2000, and 2003.

Data queries were also performed to identify individuals with lead poisoning using different BLL thresholds in units of µg/dL. In this study, we report the findings associated with the standard threshold limit recommended by the CDC, ≥10 µg/dL, hereinafter termed the elevated BLL. Additional queries were performed to identify elevated BLLs by age and sex. The children were divided into seven age groups (0–12, 13–24, 25–36, 37–48, 49–60, 61–72, and >72 months). The age-specific prevalence levels and gender-specific differences within the sample population were calculated.

To estimate the BLL prevalence rates, the number of children with elevated BLLs was divided by two denominators: (1) the sample population in the CDPH data set representing the number of children tested at least once during the year under investigation and (2) the population of all children (age zero to five years) in the study area. The former represents the BLL prevalence rate among all children who were actually screened, whereas the latter represents the BLL prevalence rate for the city’s at-risk population of children zero to five years in each census block. Both indexes were calculated at a rate of cases per 1,000. The 95 percent confidence intervals (CI) were calculated for the total number of children in the database, males, females, and the different age groups that had reported elevated BLLs.

The data were also aggregated into census tracts, community areas, and neighborhoods to create profiles of high-risk neighborhoods. The process of data aggregation at the neighborhood level to evaluate health disparities is not new. Over the last decade, several researchers have advocated this approach based on the notion that health outcomes, particularly those arising from environmental exposures, are the result of contextual factors and processes operating at the scale of whole communities and larger geographical entities (Macintyre, Maciver, and Sooman 1993; Diez-Roux 2001; Dunn and Cummins 2007). Neighborhoods have subsequently been defined and operationalized using methods that focus on the inner attributes such as sociodemographic,
infrastructural, and physical environmental characteristics; the geographical scale; and proximity of spatial units (Lebel, Pampalon, and Villeneuve 2007).

In this study, we analyzed elevated BLL prevalence rates in Chicago neighborhoods at different spatial levels. The city has many census blocks ($n = 24,691$), census block groups ($n = 2,506$), census tracts ($n = 882$), and a total of seventy-seven community areas. The community areas were aggregated further into six major neighborhoods: Northside, Northwest side, Westside, Downtown, Southside, and Far Southside (Figure 2). Consistent with the approaches used in previous applications of neighborhood health analyses, the boundaries of these six communities were preset based on local knowledge of the study area, demographics, and location of major roadways and rivers. This level of aggregation was most useful for assessing the disparate patterns of disease risk using the well-known epidemiological approach, odds ratio. Among the six neighborhoods, the Northside was designated as the reference category for comparing the other communities because of its relatively low prevalence rates of BLLs and predominantly white and affluent residential areas.

**Spatial and Geostatistical Analyses**

Geostatistical analysis, using kriging, was applied to the BLL prevalence data in an effort to account for spatial variability and identify the local anomalies in disease risk. A prior review of geostatistical theory and applications revealed a well-established tradition and empirical evidence to support the application of kriging techniques to health data. One of the first known applications was reported in Cressie (1986), and other applications have since been reported in various publications (Oliver and Webster 1990; Cressie 1991; Rushton et al. 1996; Diggle, Tawn, and Moyeed 1998; Moore and Carpenter 1999; Diggle 2000; Waller and Gotway 2004). The work of Goovaerts and coauthors, more recently, has made a huge attempt to popularize this approach using cancer data (Goovaerts and Jacquez 2004; Goovaerts 2005, 2006). The approach has successfully been applied to infer information about neighboring observations using the mean, and the resulting estimates have been determined to better reflect spatial relationships (Goovaerts et al. 2003; Goovaerts and Jacquez 2004; Oyana 2004; Guo et al. 2006).

In this study, we applied similar approaches to produce risk estimates of elevated BLLs in Chicago. The rationale for applying kriging to the BLL prevalence data was based on a number of data challenges identified during data exploration and visual screening. First, there were missing values for approximately 15 percent of the census blocks in the BLL data, suggesting the need for statistical imputation to generate unbiased risk estimates across the entire study area. Specifically, a review of the census block-level data revealed the missing values along with a large variability in the distribution of the at-risk population with a mean of 10.67 and a standard deviation of 15.1 children. The number of elevated BLL cases also varied with an overall average of 2.3 cases per census block and a standard deviation of 2.32 cases. The highest number of BLL cases (106) was observed in one of the census blocks; however, there were also several blocks with no reported cases, this being the primary cause of the missing values in the data. In evaluating the various analytical options for estimating such distributions with missing values, kriging, an optimal interpolation method with its associated Best Linear Unbiased Estimator properties, was deemed the most effective method.

Second, a review of the BLL data distribution confirmed that the raw values exhibited a significantly large positive skew (18.31; $SE = 0.034$), implying the need for data transformation prior to any advanced statistical analysis. Kriging was selected because it is a multistep procedure that allows for such transformations either prior to or during the computational process by using algorithms such as the Gaussian model to derive the empirical semivariograms and risk estimates.

Third, the spatially aggregated BLL data were plagued by the small numbers problem, with unstable prevalence levels in census blocks with small at-risk populations. A query of the data using the at-risk population mean as the threshold (i.e., fewer than ten children) led to the identification of several census blocks with small numbers. Although the highest observed at-risk population was 340 children, about a third of the census blocks (34.8 percent) had smaller population sizes ranging between one
and ten children. Past studies have noted that producing disease maps that ignore the noise and uncertainty introduced by these small numbers would likely result in unreliable risk estimates with limited applications in health policy formulation (Goovaerts et al. 2003; Goovaerts and Jacquez 2004; Goovaerts and Gebreab 2008). To overcome this problem, these scholars have suggested the use of kriging as an ideal method for smoothing the data.

So overall, given its ability to address the data problems already noted, including handling large data sets, and statistically smoothing the data while incorporating additional attributes such as the underlying spatial dependence structure, variance, and uncertainty, we were compelled to use the kriging technique to spatially predict the distribution of elevated BLL prevalence across Chicago using the data for 1997, 2000, and 2003. Previous studies faced with similar data challenges also selected this approach. Carrat and Valleron (1992), Goovaerts et al. (2003), Goovaerts and Jacquez (2004), Goovaerts (2005, 2006), and Torok et al. (1997) used kriging to handle issues of poor data distribution and small numbers, and Ribeiro et al. (1996), Rushton et al. (1996), and Kelsall and Eld (2002) used this approach to tackle missing data issues. Other classic examples are reviewed in Diggle, Tawn, and Moyeed (1998), Diggle (2000), and Waller and Gotway (2004).

The ESRI ArcGIS’s Geostatistical Analyst extension provided the tools for data exploration and analysis to resolve the problems. To complete this process, we created a centroid for each census block, derived the mean elevated BLL value for all the subjects tested in each census block, and fitted the best predictive model to estimate BLL prevalence rates for each of the census blocks. Data transformation using the natural logarithm of the prevalence rates at census block level was performed and the resulting values were kriged to generate the predicted prevalence rates. We tested for the presence of spatial autocorrelation using a variogram and the established coefficients were utilized to fit the best model. The data were then back-transformed by calculating the inverse logarithm to produce the “kriged BLL prevalence.” In investigating the residual errors, we established that our smoothed BLL prevalence maps corresponded to the original BLL data set and reflected the unbiased estimates with least error taking into account existing spatial autocorrelation.

Finally, in an effort to identify the local anomalies of disease risk, we explored the kriged BLL prevalence rates using TerraSeer’s Space-Time Intelligence Systems (STIS). This statistical package provides a variety of tools to explore and analyze spatial patterns including the local indicator of spatial association (LISA) statistic (Anselin 1995; Goovaerts and Jacquez 2004). This measure was used to clarify how closely the BLLs in one census block were related to neighboring blocks, as well as for plotting local spatial clusters of BLLs in the study region. Surface maps were generated to display the spatial distribution of elevated BLLs among children.

**Statistical and Epidemiological Analyses**

Statistical and epidemiological analyses of the BLL data were performed using the Statistical Program for Social Sciences (SPSS) Version 14.0. The epidemiological methods based on odds ratios were used to compute the odds and spatial risk of having elevated BLLs among screened children living in high-risk versus low-risk neighborhoods (at \( p \leq 0.05 \)). There is undoubtedly a long tradition and well-grounded literature supporting the modeling of disease counts and rates using this approach (Cressie 1991; Diggle 2000; Waller and Gotway 2004). Our primary interest in applying this approach, however, was to estimate the varying risks across neighborhoods using the Northside as a reference category.

Also relevant in our statistical analysis was the need to create a profile of high-risk areas based on housing age and the socioeconomic and racial and ethnic composition of the population. This was accomplished using descriptive discriminant analysis, a technique that identifies the prime indicators that reliably differentiate between two or more groups of spatial units or entities. This analysis was performed at the census tract level because nearly all of the socioeconomic and racial and ethnic data were reported at this level. Eight independent variables were used: percent whites, blacks, Latinos, Asian Americans, Native Americans, median housing age (years), median household income ($) and percentage of renter-occupied
housing. Prior to the statistical analysis, the results obtained from the LISA statistics, identifying the hot spots for elevated BLLs, were superimposed on the census geography of the city. A spatial query was then performed to identify all census tracts with identifiable hot spots for elevated BLLs. A binary variable was created to serve as the dependent categorical variable, designating tracts with BLL hot spots as 1 and those without the hot spots coded as 0. The enriched data layer consisting of the sociodemographic variables noted earlier and the risk variable was then exported into SPSS to generate a discriminant profile of risk areas in the city. The discriminant analysis was preceded by a diagnostic test to detect any potential problems of multicollinearity among the sociodemographic indicators. No significant problems were found; the highest correlation, observed among blacks and whites ($r = -0.75$), was considered to be acceptable for inclusion of both variables in the analysis.

**Results**

**Descriptive Statistics**

Approximately 13 percent (60,974) of the children in the CDPH database had elevated BLLs ($\geq 10 \mu g/dL$) at some point during the study period. The deduplicated database was made up of 50.25 percent males ($n = 235,845$) and 48.23 percent females ($n = 226,363$). Approximately 94 percent of the children were aged seventy-two months or younger; the mean age of subjects tested was 3.31 years; and the mean BLL for the children was 5.5 $\mu g/dL$, with a wide standard deviation of 4.91.

Table 1 shows the mean elevated BLL prevalence rates for the children by blood sample type. The data show a notable decline in BLL prevalence rates of approximately 74 percent over the seven-year period. The initial years of testing (1997 and 1998) showed tremendous variability (as reflected in the standard deviations), but the elevated BLL prevalence rates subsequently declined with greater consistency in the latter years. Overall, the temporal trends suggest that there are significant reductions in prevalence rates in Chicago, a pattern that is compatible with the nationwide trends.

Table 2 illustrates a more detailed assessment of elevated BLL prevalence rates by age, sex, and blood sample type for three specified years. Gender- and age-specific differences were observed across the study period. In 1997, there were slightly more boys with elevated BLLs than girls. Among the different age groups, children between zero and twelve months exhibited by far the lowest BLLs, whereas the groups with the highest prevalence levels were aged twenty-five to thirty-six months. These overall patterns were repeated in 2000 and 2003, revealing greater risks for boys and children between two and three years of age.

**Spatial Characterization of BLL Cases**

Figure 4 depicts the mean BLL prevalence rates based on the spatially aggregated data derived from the raw (unkriged) and kriged estimates in 1997, 2000, and 2003. The advantages of the kriged estimates, reflecting more stable rates, are readily evident in the comparative bars produced across the six areas. These charts also show the error bars depicting the standard deviation or level of uncertainty associated with each risk estimate. These measures confirm that there are large variations in the raw estimates obtained over the three time periods; however, those variations are substantially lower for the kriged prevalence rates. The kriged estimates also reflect better spatial relationships across the neighborhoods as evidenced by the global Moran’s $I$ values reported in Figure 4.
Table 2 Elevated blood lead levels in 1997, 2000, and 2003 by age, sex, and blood sampling type

<table>
<thead>
<tr>
<th>Year</th>
<th>Variables</th>
<th>Elevated blood lead level (≥ 10 µg/dL)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>Sex</td>
<td>Male 26.32% (13,771/52,331) 25.94–26.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female 24.35% (12,212/50,149) 23.98–24.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age (months)</td>
<td>0–12 9.62% (1,430/14,870) 9.15–10.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13–24 27.75% (4,451/16,037) 27.06–28.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>25–36 31.40% (4,636/14,766) 30.65–32.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>37–48 29.66% (5,339/18,000) 28.99–30.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>49–60 26.83% (5,157/19,224) 26.20–27.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>61–72 25.06% (3,606/14,388) 24.36–25.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type of blood test</td>
<td>Venous 25.71% (23,561/91,640) 25.43–25.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Capillary 22.51% (2,473/10,988) 21.73–23.30</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Sex</td>
<td>Male 14.89% (7,789/52,298) 14.59–15.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female 13.43% (6,723/50,047) 13.14–13.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age (months)</td>
<td>0–12 5.03% (920/18,292) 4.71–5.36</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13–24 14.57% (2,584/17,729) 14.06–15.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>25–36 19.45% (2,876/14,784) 18.82–20.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>37–48 17.29% (2,919/16,879) 16.72–17.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>49–60 15.31% (2,657/16,879) 14.78–15.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>61–72 14.78% (1,939/13,117) 14.18–15.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type of blood test</td>
<td>Venous 15.61% (13,408/85,887) 15.37–15.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Capillary 7.54% (1,270/17,308) 6.95–7.74</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>Sex</td>
<td>Male 7.36% (4,249/57,707) 7.15–7.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female 5.93% (3,291/55,484) 5.74–6.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age (months)</td>
<td>0–12 2.28% (467/20,497) 2.08–2.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13–24 7.19% (1,435/19,952) 6.84–7.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>25–36 9.25% (1,538/16,621) 8.82–9.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>37–48 8.08% (1,438/17,806) 7.68–8.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>49–60 7.12% (1,234/17,339) 6.74–7.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>61–72 6.69% (821/12,265) 6.26–7.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type of blood test</td>
<td>Venous 7.27% (6,407/88,146) 7.10–7.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Capillary 4.53% (1,291/28,491) 4.29–4.78</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Prevalence rates were calculated using the total number of children in the Chicago Department of Public Health (CDPH) database as the denominator. Prevalence rates for elevated blood lead levels were calculated using the CDPH surveillance database, 1997, 2000, and 2003 at p ≤ 0.05.

Having confirmed the effectiveness of the kriged estimates, a series of risk maps was produced for the three specified study periods, 1997, 2000, and 2003 (see Figure 5). The maps offer an accurate portrayal of the localized risks of the disease in the study area that is consistent with the temporal patterns described earlier. Geographically, the maps show that Chicago’s Westside has the highest risk of elevated BLLs. This is followed in descending order by block groups on Chicago’s Southside and Far Southside. The high-risk areas are fairly consistent over the seven-year period, yielding persistent and visually distinct hot spots for pediatric lead poisoning. A review of the blood test results obtained among children residing in these three regions (Westside, Southside, and Far Southside) confirmed the alarmingly high prevalence rates among the screened children ranging from 72 to 100 percent reported with elevated BLLs.

The visual patterns of disease risk across Chicago were corroborated further by the results from the odds ratio analysis (Table 3). For
Figure 4  Mean elevated blood lead level prevalence rates ($\geq 10 \mu g/dL$) and the respective standard deviations (depicted as error bars measuring the level of uncertainty associated with each measure). There are wide variations in unkriged prevalence rates, but these variations were reduced dramatically after kriging was applied. Moran’s I statistics for unkriged values were 0.163, 0.126, and 0.083 for 1997, 2000, and 2003, respectively, and kriged values were 0.886, 0.878, and 0.857 for 1997, 2000, and 2003, respectively.
Table 3  Odds ratios analysis of the Chicago Department of Public Health blood lead level data set for 1997, 2000, and 2003

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>CI (95%)</td>
<td>OR</td>
<td>CI (95%)</td>
<td>OR</td>
<td>CI (95%)</td>
</tr>
<tr>
<td>Westside</td>
<td>3.97</td>
<td>3.74–4.20</td>
<td>4.10</td>
<td>3.79–4.44</td>
<td>3.29</td>
<td>2.97–3.65</td>
</tr>
<tr>
<td>Southside</td>
<td>3.05</td>
<td>2.87–3.23</td>
<td>3.18</td>
<td>2.96–3.42</td>
<td>2.45</td>
<td>2.23–2.69</td>
</tr>
<tr>
<td>Far South</td>
<td>2.49</td>
<td>2.34–2.66</td>
<td>3.03</td>
<td>2.78–3.30</td>
<td>2.12</td>
<td>1.91–2.36</td>
</tr>
<tr>
<td>Downtown</td>
<td>1.76</td>
<td>1.63–1.90</td>
<td>1.56</td>
<td>1.39–1.75</td>
<td>1.43</td>
<td>1.23–1.67</td>
</tr>
<tr>
<td>Northwest</td>
<td>0.68</td>
<td>0.61–0.76</td>
<td>0.84</td>
<td>0.74–0.97</td>
<td>0.85</td>
<td>0.72–0.99</td>
</tr>
</tbody>
</table>

Notes: The odds ratios (OR) for elevated (≥ 10 µg/dL) blood lead levels (BLLs) at p ≤ 0.05. The odds of being exposed to lead concentrations were especially highest among children residing in the Westside neighborhood, followed by the Southside and Far Southside neighborhoods.

Figure 5  The spatial distribution of blood lead level prevalence rates among sampled children in 1997, 2000, and 2003. The left panel illustrates kriged maps of 1997 prevalence rates per 1,000. There are high blood lead level spatial concentrations in the Westside, Southside, and Far Southside neighborhoods. The middle panel, illustrating kriged rates for 2000, confirms a reduction in prevalence when compared to the year 1997. The right panel shows further reductions in lead prevalence rates in 2003, although pockets of high risk remain in the Westside, Southside, and Far Southside neighborhoods.
The disparities were lower in 2003, with a lesser but still significant risk for children in these neighborhoods. The children in the Downtown area were also vulnerable to higher BLLs during the study period although the odds ratios were relatively low. The results also revealed that unlike the other neighborhoods, children residing in the Northwest faced risks that were comparatively similar to their counterparts in the reference neighborhood on the Northside.

**Delineation and Profiles of High-Risk Neighborhoods**

Consistent with the previously stated objectives of this study, the geostatistical analysis...
of the pediatric lead poisoning cases yielded distinct disease clusters in Chicago. Furthermore, the analysis of the kriged rates confirmed that children were most susceptible in three areas: the Westside, Southside, and Far Southside, where the risk estimates were significantly greater than the city’s Northside. The final analytical step was to examine these areas as neighborhoods of concern based on the BLL surveillance data, the census geography characterizing the racial and ethnic and demographic profiles, and local knowledge of the communities. Delineating these areas of concern was deemed to be an important step toward the strategic planning of lead prevention programs in the city.

Figure 6 pinpoints the areas of concern based on the surface maps generated from LISA statistics. Of particular interest are the BLL hot spots, areas with high prevalence rates that are surrounded by neighboring areas with similarly high rates. A discriminant analysis was performed to compare the sociodemographic profiles of these hot spots relative to the low-risk areas (Table 4). The results suggested that there were statistically significant differences in the housing, racial, and socioeconomic composition of the geographic areas, \(\chi^2(8) = 150, \ p < 0.001\). The absolute value of the standardized function coefficient, representing the degree to which each variable contributed to the discriminant function, showed that median household income was the most relevant indicator of lead poisoning disparities in Chicago. Housing age emerged as the second most influential indicator. On average, the age of housing in the high-risk areas was slightly over fifty-five years, substantiating concerns about Chicago’s pre-1950s housing stock as a lingering source of childhood lead exposure in the city.

When examining the racial and ethnic composition of the risk areas, the structure coefficients obtained from the discriminant analysis confirmed that the locations of Latino, Asian, and black populations were moderately strong indicators of the observed disparities in elevated BLLs. These three groups were all overrepresented in the high-risk areas. The structure matrix, which reflects the correlation between the variables and the discriminating function variables, is shown in Table 4. The discriminant function is statistically significant, suggesting that there are areal differences in the socioeconomic and racial composition of the population by BLL risk zones. The function coefficients show that income and housing age are the strongest indicators, followed by the distribution of minority populations. Income and percentage whites are negatively correlated with the discriminant function.

### Table 4 Discriminant analysis of risk areas of elevated lead levels

<table>
<thead>
<tr>
<th>Variable</th>
<th>Census tracts with BLL hotspotsa</th>
<th>Census tracts without BLL hotspotsb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median household income ($)</td>
<td>$27,326</td>
<td>$41,244</td>
</tr>
<tr>
<td>Whites (%)</td>
<td>32.77</td>
<td>39.62</td>
</tr>
<tr>
<td>Blacks (%)</td>
<td>43.88</td>
<td>40.6</td>
</tr>
<tr>
<td>Native American (%)</td>
<td>0.33</td>
<td>0.32</td>
</tr>
<tr>
<td>Asians (%)</td>
<td>5.60</td>
<td>3.45</td>
</tr>
<tr>
<td>Latinos (%)</td>
<td>26.19</td>
<td>21.34</td>
</tr>
<tr>
<td>Median housing age (years)</td>
<td>55.27</td>
<td>50.05</td>
</tr>
<tr>
<td>Renter occupied housing (%)</td>
<td>55.59</td>
<td>47.31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients variable</th>
<th>Standardized function coefficients</th>
<th>Correlation coefficients with discrimination function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median household income ($)</td>
<td>-0.768</td>
<td>-0.801</td>
</tr>
<tr>
<td>Whites (%)</td>
<td>0.001</td>
<td>-0.203</td>
</tr>
<tr>
<td>Blacks (%)</td>
<td>0.317</td>
<td>0.075</td>
</tr>
<tr>
<td>Native American (%)</td>
<td>-0.140</td>
<td>0.007</td>
</tr>
<tr>
<td>Asians (%)</td>
<td>0.359</td>
<td>0.228</td>
</tr>
<tr>
<td>Latinos (%)</td>
<td>0.406</td>
<td>0.136</td>
</tr>
<tr>
<td>Median housing age (years)</td>
<td>0.459</td>
<td>0.496</td>
</tr>
<tr>
<td>Renter occupied housing (%)</td>
<td>-0.005</td>
<td>0.395</td>
</tr>
</tbody>
</table>

Notes: BLL = blood lead level. The discriminant function is statistically significant, suggesting that there are areal differences in the socioeconomic and racial composition of the population by BLL risk zones. The function coefficients show that income and housing age are the strongest indicators, followed by the distribution of minority populations. Income and percentage whites are negatively correlated with the discriminant function.

\(a n = 218\).

\(b n = 648\). For (B) Wilks’s \(\Lambda = 0.840; \chi^2 = 150.14 (8df), \ P \leq .001, \) valid \(N = 862\).
function, showed a negative relationship between the function and two variables, median household income and percentage whites. All of the other variables were positively associated with the function. These results, overall, confirmed that statistically, the high-risk areas for elevated BLLs were most associated with low-income and non-white populations and a high proportion of old residential units occupied mostly by renters. Using these results, the profiles of three neighborhoods identified as major areas of concern were generated. A brief summary of the neighborhood characteristics is provided next.

The first area of concern, the Westside community, had the highest prevalence rates of elevated BLLs in all categories. Analysis of the census geography confirmed that this area is populated largely by African Americans and Hispanics within the census blocks and in scattered pockets adjacent to this community. The area is highly urbanized, with female-headed households and many children (U.S. Census Bureau 2000). It is also an impoverished community with poor housing conditions and all of the reported risks for childhood lead toxicity.

The second area of concern, the Southside, had the second highest prevalence rates of elevated BLLs among children. This community is populated by minorities, mostly blacks in the east and Latinos in the west. Several of the neighborhoods are socioeconomically disadvantaged, with a moderately large stock of pre-1950s housing and all of the other confirmed risks for childhood lead toxicity.

The third area of concern, the Far Southside, had the third highest disease rates in comparison with other communities. The community’s racial composition is diverse, although there are specific locations with significant concentrations of minorities. One of those locations was identified as the Altgeld Gardens, a neighborhood that hosts one of the largest rental housing projects with more than 10,000 people. Concerns about lead pollution and other hazardous chemicals in this area are well documented (Cohen 1992; U.S. Environmental Protection Agency 1998; Pellow 2002; Binns et al. 2004), though previous environmental tests found no apparent public health hazards (U.S. Environmental Protection Agency 1998). Our findings, however, including those of other recent studies (Lanphear, Matte, et al. 1998; Shinn et al. 2000; Binns, Kim, and Campbell 2001; Binns et al. 2004), show unusually high BLL prevalence rates in the area. It is, therefore, included among the three neighborhoods that deservedly require immediate intervention for lead poisoning.

Summary and Research Implications

Using a large-scale surveillance database, this study has shown that pediatric lead poisoning, although declining in prevalence rates, still poses a serious risk, particularly in minority and socioeconomically distressed neighborhoods in Chicago. A comprehensive analysis of the data based on a rigorous set of statistical and geostatistical approaches revealed four main findings: (1) the prevalence rates of elevated BLL decreased significantly for both genders and all six age groups during the study period (1997–2003); (2) children between the ages of twenty-five and thirty-six months consistently had the highest levels of elevated BLLs; (3) geographically, the Westside, Southside, and Far Southside neighborhoods consistently had the highest rates of elevated BLLs; and (4) the high-risk areas were overrepresented by minority and low-income families residing in renter-occupied and older housing units.

The study’s revelation that elevated lead levels declined significantly in Chicago (74 percent) over the years is consistent with studies conducted nationwide over the last two decades. Since regulations were put in place to ban lead-based paint and other by-products and promote the early identification of the disease through mandatory screening of young children, BLL prevalence rates have declined steadily each year.

Also significant was the persistent risk of lead exposure observed in some Chicago neighborhoods despite the dramatic declines in BLLs citywide. This study revealed that three neighborhoods (notably the Westside, Southside, and Far Southside) not only had high prevalence rates but that children were consistently at risk over the seven-year period. The most important risk factors in these neighborhoods were related to housing age and poverty among underrepresented populations. These findings, suggesting the ongoing disparities by race and class, are compatible with earlier reports.
Among the individual risk factors associated with elevated BLLs, the age of the children, particularly those between twenty-five and thirty-six months, who had the highest rates, was a consistent risk factor. The youngest children, aged zero to twelve months, had the lowest prevalence rates. These findings reflect the limited exposure to lead contaminants among newborns. As children get older and interact with their physical environment more readily, however, their hand-to-mouth activity increases, which amplifies the likelihood of ingesting lead-contaminated substances. It is therefore recommended, as the state and local health department agencies have done, that children continue to receive annual screenings until they reach the age of six.

One of the strengths of this study was demonstrating the use of geostatistical techniques in targeting high-risk neighborhoods for pediatric lead prevention programs. The use of kriging allowed for spatial prediction and the generation of optimal estimates of BLL prevalence levels. These estimates, in turn, led to the detection of persistent clusters of the disease risk across time and space. Extracting the spatial estimates for statistical validation in SPSS enabled the comparison of disease risks, using odds ratios, and finally delineating the racial and economic profile of the neighborhoods of concern. It is highly recommended that children in these high-risk neighborhoods be screened regularly to identify those with elevated BLLs.

It is important to note that kriging has its share of methodological pitfalls despite the fact that, overall, it produced better local estimates of BLLs and smaller estimation variances of the disease. As with most interpolative methods, the technique suffers from potential problems of underestimation or overestimation of disease rates. Ordinary kriging, in particular, tends to produce overly smoothed estimates that might not be entirely representative of the disease rates in the community. For example, in this study, the continuous surface risk maps produced from the kriged estimates might incorrectly imply that blood lead prevalence levels are continuously distributed across the city. The smoothing of the data potentially reduces the sharp distinctions in disease distribution, a pattern that might be particularly useful when tracking disease hot spots. In this study, we were able to address this problem by comparing both the outcomes produced by unkriged and kriged procedures with the LISA statistics and then using the TerraSeer program to detect the BLL hot spots. In pursuing both estimates we succeeded in producing a robust disease model showing consistent and persistent patterns and trends of BLLs in the study area.

There are other problems uncovered in this study that are worthy of further exploration. One of these has to do with conducting the health analysis at multiple spatial scales. This was done to maximize the use of the BLL data aggregated initially at the census block level, even though the ancillary data on race and class were available at the tract level. The implications, particularly in terms of the modifiable areal unit problem, are worthy of further investigation. Another issue that was not addressed in this article is the mobility of children over the study period. Given the unique identifier of each subject in the database, it might be worthwhile to track the residential patterns and mobility of the children in an effort to pinpoint the geographic source of lead exposure. Beyond these analytical challenges, additional research is required to (1) evaluate the long-term health effects—relating elevated BLLs to developmental problems—such as learning and behavioral disorders and (2) conduct in-depth, population-based, and cohort studies in the designated areas of concern. Analysis of child literacy levels, soil and water quality data, and proximity to major roadways would assist in establishing the lead pathways and potential health outcomes among children in disadvantaged communities.

### Literature Cited


TONNY J. OYANA is an associate professor in the Department of Geography at the Southern Illinois University, Carbondale, IL 62901. E-mail: tjoyana@siu.edu. His research interests include GISystems, GIScience, cartographic design, geography and algorithms, pattern recognition, environmental health and exposure, and spatial analysis.

FLORENCE M. MARGAI is an associate professor in the Department of Geography at Binghamton University, Binghamton, NY 13902. E-mail: margai@binghamton.edu. Her research interests include pediatric health geographies, environmental health hazards, health disparities, environmental justice, spatial analysis, and geostatistics.