Original Article

Residence Proximity to Benzene Release Sites Is Associated With Increased Incidence of Non-Hodgkin Lymphoma

Catherine Bulka, MPH[1]; Loretta J. Nastoupil, MD[2]; William McClellan, MD, MPH[1]; Alexander Ambinder, BA[2]; Adrienne Phillips, MD[3]; Kevin Ward, PhD, MPH[1,4]; A. Rana Bayakly, MPH[5]; Jeffrey M. Switchenko, PhD[2,6]; Lance Waller, PhD[6]; and Christopher R. Flowers, MD, MS[2]

BACKGROUND: An increased risk of non-Hodgkin lymphoma (NHL) has been observed among individuals with occupational exposure to benzene, but the risk among those living near benzene release sites has not been well described. **METHODS:** To investigate the spatial patterns of NHL incidence and the association between NHL incidence and distance to benzene release sites, the authors linked and geocoded data on benzene release sites in Georgia from 1988 to 1998 using the Environmental Protection Agency's (EPA) Toxics Release Inventory (TRI), census tract level population statistics, and NHL incidence from the Georgia Comprehensive Cancer Registry (GCCR) from 1999 to 2008. Standardized incidence ratios were mapped by census tract, and a Poisson regression was performed on NHL and NHL subtype incidence data using the mean distance between the tract centroids and release sites as markers of exposure. Cluster analyses were conducted at the global, local, and focal levels. **RESULTS:** Poisson regression indicated that, for every mile the average distance to benzene release sites increased, there was an expected 0.31% decrease in the risk of NHL. Similar results were observed for all NHL subtypes analyzed. Clusters of NHL were spatially associated with benzene release sites located in metropolitan areas, but not with release sites in other areas of the state. **CONCLUSIONS:** NHL incidence was significantly higher in census tracts that were closer, on average, to benzene release sites. Additional studies are needed to examine spatial patterns of NHL incidence in other geographic regions and interactions between benzene and other exposures. *Cancer* 2013;119:3309-17. © 2013 American Cancer Society.

KEYWORDS: lymphoma; non-Hodgkin; geographic information systems; benzene; surveillance.

INTRODUCTION

Since the 1970s, the incidence of non-Hodgkin lymphoma (NHL) has increased by 3% to 4% annually[1] driven by modifications of the lymphoma classification system, improved techniques for diagnosis and reporting hematologic malignancies, and the human immunodeficiency virus (HIV) epidemic. However, these factors account for approximately 50% of the additional cases of NHL.[2,3] The increase in NHL diagnoses also appears to trail expanded industrial production in the United States, suggesting that occupational chemical exposures are risk factors for NHL.[4]

Among carcinogenic occupational chemicals, benzene consistently has been linked to hematologic cancers.[5] Numerous studies in human populations support a link between occupational benzene exposure and NHL, although some controversy remains.[6-8] Biologically, benzene exposure produces chromosomal aberrations and genetic changes, and toxic effects of benzene exposure can occur at air levels of 1 part per million or less, suggesting that even low levels of benzene exposure can be harmful.[9] Additional research is necessary to investigate associations between benzene exposure and NHL in residential settings, because urban and rural populations are exposed to benzene through release into ambient air and surface waters. The purpose of the current study was to examine relations between the incidence of NHL at the census tract level and benzene release sites to determine whether NHL incidence varies with distance from these sites.

MATERIALS AND METHODS

We linked and geocoded data on benzene release sites in Georgia from 1988 to 1998 from the Environmental Protection Agency's (EPA) Toxics Release Inventory (TRI), NHL incidence from the Georgia Comprehensive Cancer Registry (GCCR) (1999-2008), and population and demographic data from the 2000 US Census. This study was approved by the Emory University Institutional Review Board, the Winship Cancer Institute Clinical Research Committee, and the Georgia Department of Public Health Institutional Review Board. Because individual patients were not contacted and unmasked data were handled according to Health Insurance Portability and Accountability Act of 1996 privacy standards, individual informed consent was not required for this study.

Benzene Release Site Data

The EPA requires that facilities meeting thresholds defined by Section 313 of Emergency Planning and Community Right-to-Know Act annually report their disposal or releases for listed toxic chemicals. Launched in 1987, the TRI captures release information about certain chemicals, including the quantities, media type, and geographic coordinates of the releases from sources like manufacturing facilities, service business, and federal facilities.[10] Between 1988 and 1998, 19 facilities in Georgia reported benzene releases. Total releases, calculated as the sum of fugitive air releases, stack air releases, and surface water discharges, between 1988 and 1998 for each site ranged from 52 to 3,830,097 pounds of benzene. Some release sites reported releases for only 1 year, whereas others reported benzene releases for multiple years during this period.

Census Tract Population and Demographic Data

Census tracts are subdivisions of counties with an average population of 4000. At the time of the 2000 Census, there were 1618 census tracts within the state of Georgia. Population and demographic data were available for 1616 of these tracts; and data on sex, age, and race were obtained from Summary File 1.[11] Because accurate yearly population values were not available for each census tract, we used 2000 census data and census tract boundaries to geocoded GCCR data and calculate standardized incidence ratios (SIRs) for each census tract. We used census data on the median year moved into residence (MYMI), which was derived from a sample of individuals residing within a census tract, to provide information on the length of time residents were in their current home and on their residential exposure to benzene release sites.

Non-Hodgkin Lymphoma Data

GCCR provided data for all 12,716 incident NHL cases among adults aged ≥20 years residing in Georgia at the time of diagnosis for the period from 1999 to 2008. We previously examined similar population-level incidence data for NHL and Hodgkin lymphoma subtypes in the United States.[12-16] To standardize the NHL incidence rates from Georgia to national NHL incidence rates, SEER/Stat software (version 7.05; Surveillance, Epidemiology, and End Results [SEER] Program; National Cancer Institute, Bethesda, Md) was used to access the SEER 13 registries database.[17] Age-specific, sex-specific, and race-specific crude incidence rates were obtained for the period from 1999 to 2008 for NHL and for NHL subtypes to standardize Georgia incidence data by age, sex, and race. Incident cases were separated into B-cell NHL (NHL-B), T-cell NHL (NHL-T), diffuse large B-cell lymphoma (DLBCL), and follicular lymphoma (FL) using *International Classification of Diseases for Oncology* third edition codes based on the proposed World Health Organization-based nested classification of malignant lymphoid neoplasms for epidemiologic research from InterLymph.[18] Patients who were missing information on age, sex, or race were excluded from all analyses. We used 4 major race categories (white, black, American Indian/Alaska Native, and Asian or Pacific Islander); patients whose race was categorized as "other" or "unknown" were excluded. Patients who were geocoded successfully were aggregated to the census tract level.[19] SIRs were calculated for each census tract by dividing the number of observed cases by expected cases.

Geographic Data and Spatial Analyses

We used geographic information system data through ArcGIS 10 (ESRI, Redlands, Calif) to examine the spatial distribution of benzene release sites and SIRs by census tract. Census tract shape files were obtained from the US Census Bureau's 2000 TIGER/Line files.[20] The geographic information system was used to calculate the distances between the 1616 census tract centroids to each of the 19 benzene release sites. Descriptive spatial analysis was performed using ArcGIS. Choropleth maps were created to depict the NHL SIRs and the NHL subtype SIRs by census tract. Locations of benzene release sites were laid over the SIR maps. Because sparse data can pose statistical problems in small area analyses, spatial smoothing was performed to allow census tracts to borrow strength from spatial neighbors and produce more stable estimates of SIRs while maintaining geographic precision. Spatial empirical Bayes (SEB) smoothing was performed on the SIR values using GeoDa 1.01 (Luc Anselin, Tempe,

Ariz).[21] We defined neighbors as census tracts with common borders or vertices and constructed choropleth maps of smoothed SIRs for NHL and for NHL subtypes.

To assess spatial correlation of SIRs, we conducted global, local, and focal spatial analyses. A global measure of spatial autocorrelation, the global Moran I, was calculated for SIR patterns of NHL and NHL subtypes using GeoDa. It is important to note that smoothing will induce additional spatial correlation; thus, we expected to observe higher and more statistically significant Moran I values for the SEB-smoothed SIRs. To measure spatial autocorrelation at the local scale, a local Moran I, also termed Local Indicators of Spatial Autocorrelation (LISA),[21] was calculated for SIR patterns of NHL and NHL subtypes using GeoDa. LISA cluster maps were created to identify the locations of any significant clusters of SIRs. Both global and local spatial statistics based significance on 999 Monte Carlo simulations with neighbors defined by shared boundaries and vertices.[21]

We used indirectly standardized incidence rates to test for focal clustering, which we calculated by multiplying the crude incidence rate by the nonsmoothed SIR, to test the hypothesis that there was no spatial clustering near the benzene release sites (the focus). The Lawson-Waller test was used to individually assess each of the 19 benzene release sites for focal clustering of NHL.[21] 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.[21] The standard normal distribution was used to estimate upper-tail P values in ClusterSeer 2.3 (BioMedware, Ann Arbor, Mich). Because the Lawson-Waller test was conducted 19 times (once for each release site), we adjusted our significance level using the Bonferroni correction method.[22]

Poisson regression models were constructed under the assumption that the number of observed incident cases for each census tract had a Poisson distribution that depended on 1) the number of expected cases for that census tract based on its age, sex, and race demographics and 2) the explanatory variable of mean distance from benzene release site. The median year moved into residence was also assessed as a potential confounder and/or effect modifier. The models were estimated using the maximum likelihood method. Models were fitted for each subtype of NHL in addition to the total NHL cases model using SAS version 9.3 (SAS Institute, Inc., Cary, NC).

RESULTS

Of the 12,716 incident cases of NHL in our original data set, 11,355 cases were successfully geocoded to the census

TABLE 1. Study Population Demographics

Variable	Cases, N = 11,323[a]		Georgia Population, N = 5,582,707[b]	
	No.	%	No.	%
Age group, y				
20–59	4585	40.5	4,521,891	81.0
≥60	6738	59.5	1,060,816	19.0
Sex				
Women	5274	46.6	2,907,559	52.1
Men	6049	53.4	2,675,148	47.9
Race				
White	8991	79.4	3,922,068	70.3
Black	2208	19.5	1,521,316	27.3
American Indian/Alaskan Native	4	0.0	15,326	0.3
Asian or Pacific Islander	120	1.1	123,997	2.2
NHL subtype				
B-cell NHL	8925	78.8		
Diffuse large B-cell lymphoma	3851	34.0		
Follicular lymphoma	2171	19.2		
T-cell NHL	1489	13.2		
Median year moved into residence			1989	
Mean distance from benzene release sites: Mean±SD, miles			99.5±55.2	

Abbreviations: NHL, non-Hodgkin lymphoma; SD: standard deviation.
[a] Total number of NHL cases in Georgia from 1999 to 2008.
[b] This population was based on 1616 census tracts in Georgia from the 2000 US Census.

tract level. Among these, 11,323 had age, sex, and race information available and could be classified into the US census race categories of "white," "black," "American Indian/Alaska Native," or "Asian or Pacific Islander" (Table 1). We mapped SIRs and SEB-smoothed SIRs for NHL, NHL-B, NHL-T, DLBCL, and FL with location data for each benzene release site (Fig. 1). Two census tracts were not included because of a lack of demographic data from the US Census Bureau. Fourteen of the 19 benzene release sites were in the greater metropolitan Atlanta area.

Spatial Analyses

The global Moran I values were greater than 0 and were significant at the $\alpha = .05$ level for the NHL, NHL-B, NHL-T, and DLBCL raw SIRs, demonstrating positive spatial autocorrelation (Table 2). This indicates that census tracts with similar incidence rates (high or low) were clustered. When spatial smoothing was performed to allow census tracts to borrow strength from spatial neighbors, there was significant positive spatial autocorrelation for the SEB-smoothed SIRs for NHL and all subtypes, which was expected because the smoothing process induced additional spatial correlation. LISA cluster maps of SEB-smoothed SIRs (Fig. 2) illustrate the locations of
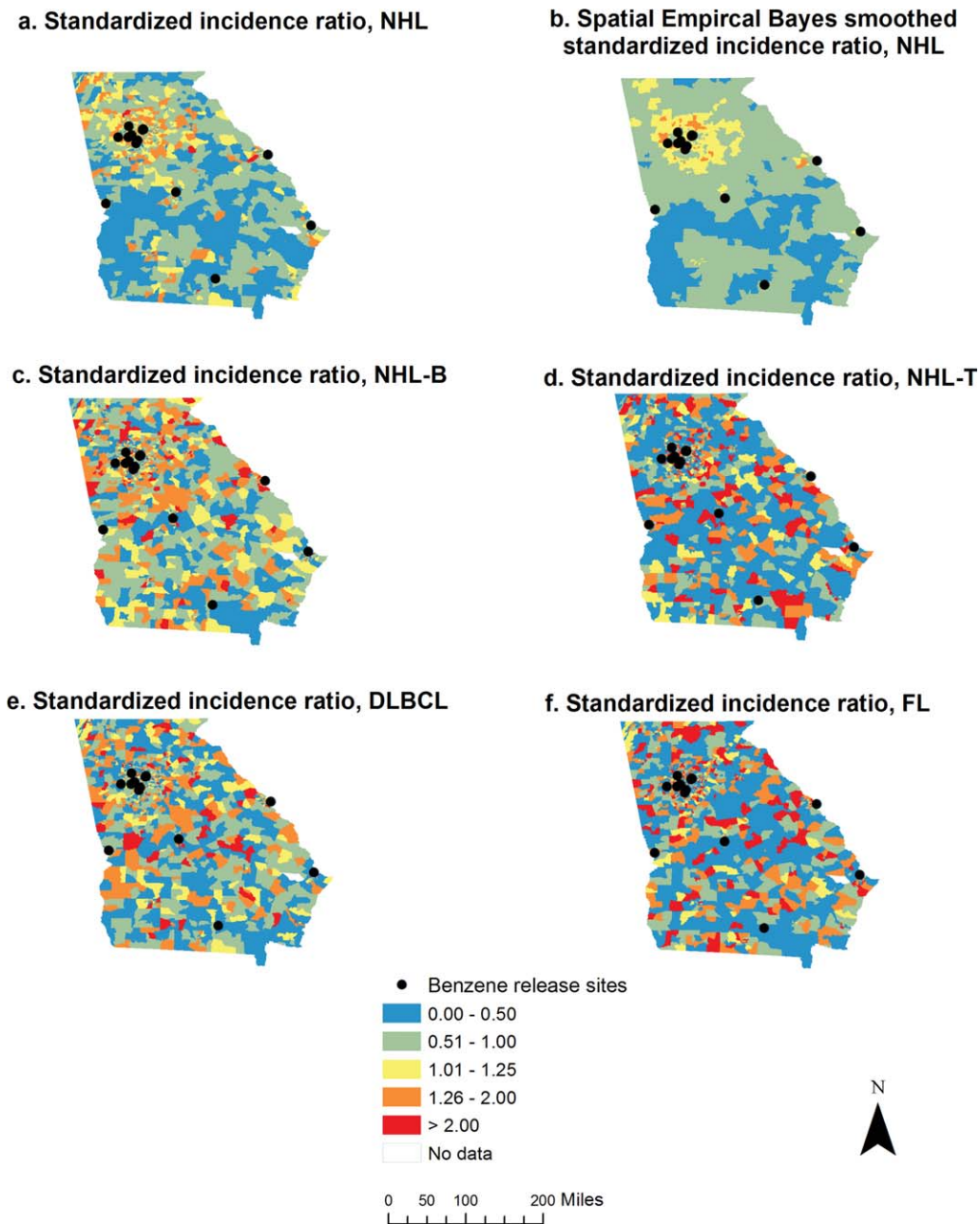
Figure 1. In these standardized incidence ratio (SIR) maps of non-Hodgkin lymphoma (NHL) among adults in Georgia from 1999 to 2008, blue represents an SIR of 0.00 to 0.50, green represents an SIR of 0.51 to 1.00, yellow represents an SIR from 1.01 to 1.25, orange represents an SIR from 1.26 to 2.00, and red represents an SIR >2.00. Maps show (a) SIRs for NHL, (b) spatial empirical Bayes-smoothed SIRs for NHL, (c) SIRs for B-cell NHL (NHL-B), (d) SIRs for T-cell NHL (NHL-T), (e) SIRs for diffuse large B-cell lymphoma (DLBCL), and (f) SIRs for follicular lymphoma (FL).

"hot-spots" (high-high clusters) and "cold-spots" (low-low clusters). High-high areas indicate significant clustering of census tracts with high SIR values surrounded by tracts that also have high SIR values, whereas low-low areas indicate significant clustering of census tracts with low SIR values surrounded by tracts that also have low SIR values. Low-high areas indicate census tracts with low SIR values surrounded by tracts with high SIR values, and high-low areas indicate census tracts with high SIR values surrounded by tracts with low SIR values. Clustering of high SIRs appears to be located in the metropolitan Atlanta area for NHL and for NHL subtypes, whereas clustering of low SIRs appears to be mostly in the southern region of the state.

For 15 of the 19 benzene release sites, the Lawson-Waller test was statistically significant at the α = .0026 level (using Bonferroni correction for 19 comparisons) (Table 3). All benzene release sites located within the metropolitan Atlanta area had significant focal clustering, with greater incidence rates observed near these sites than expected. Incidence rates were lower than expected surrounding sites 6, 8, 9, and 10. Only 1 benzene release site associated with increased NHL incidence was located outside of the metropolitan Atlanta area;

this site was located near Savannah, Georgia. No statistically significant Poisson coefficients emerged for the interaction term between mean distance to benzene release site and MYMI, so this variable was dropped. In addition, comparing models with and without MYMI suggested that it was not a confounder. A 0.31% decrease in NHL risk is expected for every mile the average distance to benzene release sites increases. Results of the Poisson regression analyses are displayed in Table 4.

DISCUSSION

We examined residential proximity to benzene release sites and risk of NHL in Georgia using census tract level data. Our Poisson regression results suggest an approximate 0.3% decrease in NHL risk is expected for a given population in a census tract when increasing the average distance to a benzene release site by 1 mile. Among NHL subtypes, mean distance had the least effect on FL incidence, but cases were still significantly more likely to occur in census tracts that were closer, on average, to benzene release sites. It is unclear why the effect was weakest on FL incidence, but recent research on FL suggests that

TABLE 2. Global Measures of Spatial Autocorrelation

Measure	NHL	NHL-B	NHL-T	DLBCL	FL
Nonsmoothed SIR					
Moran I	0.2009	0.0628	0.0430	0.0179	−0.0011
P	.0010	.0010	.0020	.1090	.5150
SEB-smoothed SIR					
Moran I	0.8317	0.6377	0.5879	0.6403	0.5938
P	.0010	.0010	.0010	.0010	.0010

Abbreviations: DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; NHL, non-Hodgkin lymphoma; NHL-B, B-cell non-Hodgkin lymphoma; NHL-T, T-cell non-Hodgkin lymphoma; SEB, Spatial Empirical Bayes; SIR, standardized incidence ratio.
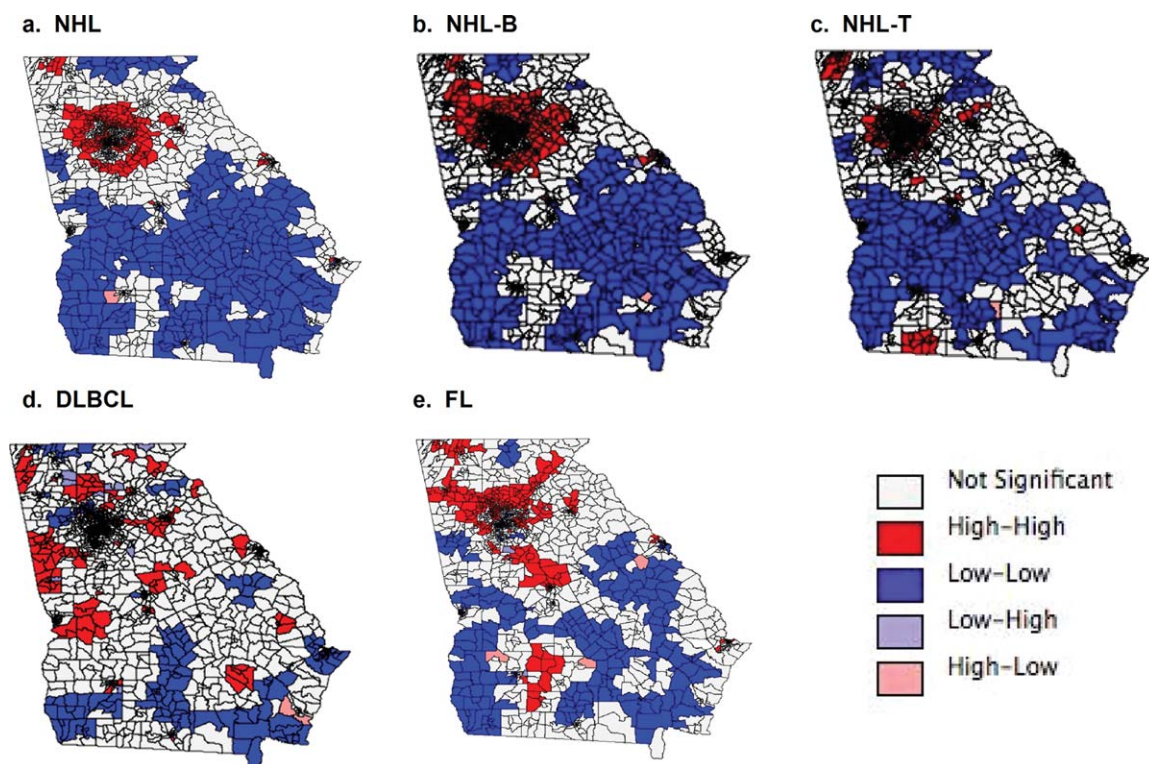

Figure 2. Local Indicators of Spatial Autocorrelation cluster maps of spatial empirical Bayes-smoothed standardized incidence ratios show (a) non-Hodgkin lymphoma (NHL), (b) B-cell NHL (NHL-B), (c) T-cell NHL (NHL-T), (d) diffuse large B-cell lymphoma (DLBCL), and (e) follicular lymphoma (FL).

genetics and lifestyle are more predictive of FL than environmental factors and that these genetic and lifestyle risk factors are associated more with FL than with other NHL subtypes.[23,24] Although the overall incidence of NHL in Georgia is low (17.4 per 100,000 population per year), NHL incidence was significantly higher in tracts that were closer to benzene release sites. 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, whereas other urban areas, namely Augusta and Savannah, were 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. These results suggest that focal clustering is present around benzene release sites located in the metropolitan Atlanta, but this phenomenon was not observed in other areas of the state. An alternative hypothesis that could explain the results observed from our cluster analyses is that populations living in urban areas are exposed to some unmeasured factor or an unaccounted for confounder. Other sources of benzene exposure associated with urban areas include mainstream smoke from cigarettes and auto exhaust.[25] However, neither of these have been consistently associated with an increased risk of NHL.[26] In addition, pesticides, which also have been associated with increased risk of NHL, are found more commonly in rural areas.[27]

Strengths of this study include the use of publicly available data from EPA and the US Census Bureau. Spatial smoothing of NHL incidence data addressed differences in census tract population size that could result in variance instability and spurious outliers. By standardizing the data indirectly, we eliminated the effects of age, sex, and race on NHL incidence. Our knowledge, this study is the first to examine the relation between passive benzene exposure and the incidence of NHL at the state population level. One limitation of this study is the use of aggregated data, because our findings may not hold true at the individual level. Furthermore, the presence of chemical releases into the environment is not sufficient to determine personal exposure or to calculate potential risks to human health.

A potential weakness of this study was the lack of quantitative exposures and temporal analyses. A nationwide study revealed that amounts of benzene emissions may be as much as 5 times higher than reported by the TRI.[28] Thus, although data on the amounts of benzene released from facilities in Georgia between 1988 and 1998 were available, these were not incorporated into the

TABLE 3. Lawson-Waller Score Test for Focal Clustering

Benzene Site	Score	P
1. Metropolitan Atlanta	13.1831	<.0001[a]
2. Metropolitan Atlanta	18.0533	<.0001[a]
3. Metropolitan Atlanta	36.3879	<.0001[a]
4. Metropolitan Atlanta	14.9471	<.0001[a]
5. Metropolitan Atlanta	15.6043	<.0001[a]
6. Homerville	−23.6137	1.0000
7. Metropolitan Atlanta	15.8019	<.0001[a]
8. Macon	−3.8011	.9999
9. Columbus	−3.4838	.9999
10. Augusta	−6.7270	1.0000
11. Metropolitan Atlanta	26.9324	<.0001[a]
12. Metropolitan Atlanta	20.8564	<.0001[a]
13. Metropolitan Atlanta	14.9068	<.0001[a]
14. Metropolitan Atlanta	15.6043	<.0001[a]
15. Metropolitan Atlanta	15.0695	<.0001[a]
16. Savannah	9.2829	<.0001[a]
17. Metropolitan Atlanta	29.9130	<.0001[a]
18. Metropolitan Atlanta	37.9698	<.0001[a]
19. Metropolitan Atlanta	18.4694	<.0001[a]

[a] These P values are significant at $\alpha = .0026$ (Bonferroni correction for multiple comparisons).

TABLE 4. Poisson Regression Results

Variable	Model 1: NHL	Model 2: NHL-B	Model 3: NHL-T	Model 4: DLBCL	Model 5: FL
Intercept	0.1407	0.3453	0.0762	2.4831	0.1460
Explanatory variables					
Mean distance from benzene release site β	−0.0031	−0.0035	−0.0032	−0.0032	−0.0020
Standard error	0.0002	0.0002	0.0006	0.0004	0.0003
Wald chi-square statistic	227.27	223.85	31.25	80.09	55.72
P	< .0001	< .0001	< .0001	< .0001	< .0001
Model fit					
Deviance, df = 1614	2599.96	2913.62	1843.55	2118.79	8964.21
Deviance/df	1.61	1.81	1.14	1.31	5.55

Abbreviations: df, degrees of freedom; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; NHL, non-Hodgkin lymphoma; NHL-B, B-cell non-Hodgkin lymphoma; NHL-T, T-cell non-Hodgkin lymphoma.

models because of concerns about reliability. However, we attempted to control for time and latency by only including benzene release sites from 1988 to 1998 and by using incidence and population data from 1999 to 2008 as well as assessing MYMI as a potential confounder and effect modifier. The average value for MYMI for all census tracts used in this investigation was 1989, which was during the period of potential benzene exposure (1988-1998) and allowed for a theoretical period of 10 years to develop NHL, because our incidence data were for the years 1999 to 2008. Changes in the population of Georgia between 1999 and 2008 and the use of 2000 US Census data as denominators for incidence rates also may have introduced bias into this study. In addition, any unmeasured confounding variables may have been spatially correlated, which could have biased our Poisson regression coefficients.

Associations between exposure to benzene and hematologic effects were observed as early as the 1940s.[8] In a 35-year cohort study of rubber workers from this period, significant decreases in white blood cell and erythrocyte counts were observed among workers exposed to benzene. More recent occupational studies have demonstrated similar associations.[5,7] In China, workers exposed to benzene for $\geq$10 years were 4.2 times more likely (95% confidence interval, 1.1-15.9 times more likely) to develop NHL.[6] The development of NHL has been linked most strongly to benzene exposures that occurred >10 years before diagnosis, suggesting that benzene-related NHL may be associated with longer induction periods. Often, it is difficult to discern the specific role benzene plays in the elevation of NHL risk, because workers exposed to benzene in an occupational setting tend to 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, 40 studies (93%) indicated an increase in NHL risk, and 23 studies (53%) demonstrated statistically significant associations between NHL risk and benzene exposure.[29] Additional meta-analyses have been supportive of this finding,[30] whereas other meta-analyses have reached different conclusions.[31] For example, a recent meta-analysis by Kane and Newton indicated no effect of occupational benzene exposure on NHL risk, regardless of the amount or duration of benzene exposure.[32] One explanation for differences in these findings is that NHL represents a heterogeneous group of diseases, and benzene exposure may be associated more strongly with some variants than others.[30] The Kane and Newton meta-analysis

demonstrated no association between occupational benzene exposure and NHL, DLBCL, or FL; whereas our study demonstrated associations between passive benzene exposure and NHL and NHL subtypes. The latency period before NHL development also may explain why some studies failed to observe an association with NHL. Future studies can examine temporal relations between passive benzene exposure and the incidence of NHL.

Our study contributes to a growing body of literature that has identified the association between NHL and organic solvents like benzene. Previous studies of nonoccupational exposure and NHL have been limited by their reliance on subject recall,[33] whereas more recent studies have used geographic approximations, such as zip code or municipality centroids, to conduct their analyses.[34-36] Johnson et al observed an increased risk of NHL in individuals living in proximity to copper smelters and sulfite pulp mills, whereas Ramis et al reported an association between NHL and proximity to paper industries. Another study indicated an increased risk of NHL in patients who lived in wards surrounding a municipal solid waste incinerator that was known to exude dioxins.[37]

A recent study by Wheeler et al used case residential histories to examine spatiotemporal clustering of NHL in Iowa, Los Angeles, Detroit, and Seattle. Despite adjusting for polychlorinated biphenyl (PCB) exposure and genetic polymorphisms (known risk factors for developing NHL), spatiotemporal clusters of high NHL incidence remained unexplained.[38] Other studies have reported associations between NHL and proximity to petroleum refineries and toxic industrial waste sites.[39,40] De Roos et al conducted a case-control study of industrial facilities and NHL in 2010 that used geographic positioning systems to identify the exact coordinates of residence for 99% of their 1548 participants.[4] Their study was unique in its ability to determine risk estimates for individual NHL subtypes and to assess trends with increasing proximity to various industrial facilities. Although those authors did not study benzene exposure specifically, their study indicated that individuals living within 2 miles of petroleum refineries (which are known to contain benzene) for >10 years had a 90% increased risk of NHL (odds ratio, 1.9; 95% confidence interval, 1.0-3.6; P_{trend} = .04). Continued improvements in the geographic information system analysis techniques used to study the associations between benzene and NHL will contribute to an improved understanding of NHL etiology.

This study identified a significant association between distance to benzene release sites and NHL incidence at the census tract level. Future studies should examine this association using individual-level geographic data, temporal trends, and dose-response relations between benzene exposure and NHL. Additional future directions for this research will evaluate other potential sources of benzene exposure, other environmental exposures, and genetic and clinical factors that may explain the higher incidence of NHL observed in certain areas. These studies investigating the etiology of NHL are critical to identify and enact public health policies that may decrease or prevent NHL.

FUNDING SOURCES

CONFLICT OF INTEREST DISCLOSURES

The authors had no disclosures related to the content of this manuscript.

REFERENCES

1. Muller AM, Ihorst G, Mertelsmann R, Engelhardt M. Epidemiology of non-Hodgkin's lymphoma (NHL): trends, geographic distribution, and etiology. *Ann Hematol*. 2005;84:1-12.
2. Hartge P, Devesa SS. Quantification of the impact of known risk factors on time trends in non-Hodgkin's lymphoma incidence. *Cancer Res*. 1992;52(19 suppl):5566s-5569s.
3. Flowers CR, Sinha R, Vose JM. Improving outcomes for patients with diffuse large B-cell lymphoma. *CA Cancer J Clin*. 2010;60:393-408.
4. De Roos AJ, Davis S, Colt JS, et al. Residential proximity to industrial facilities and risk of non-Hodgkin lymphoma. *Environ Res*. 2010;110:70-78.
5. Smith MT. Advances in understanding benzene health effects and susceptibility. *Annu Rev Public Health*. 2010;31:133-148.
6. Hayes RB, Yin SN, Dosemeci M, et al. Benzene and the dose-related incidence of hematologic neoplasms in China. Chinese Academy of Preventive Medicine—National Cancer Institute Benzene Study Group. *J Natl Cancer Inst*. 1997;89:1065-1071.
7. Rinsky RA, Hornung RW, Silver SR, Tseng CY. Benzene exposure and hematopoietic mortality: a long-term epidemiologic risk assessment. *Am J Ind Med*. 2002;42:474-480.
8. Kipen HM, Cody RP, Crump KS, Allen BC, Goldstein BD. Hematologic effects of benzene: a thirty-5 year longitudinal study of rubber workers. *Toxicol Ind Health*. 1988;4:411-430.
9. Bollati V, Baccarelli A, Hou L, et al. Changes in DNA methylation patterns in subjects exposed to low-dose benzene. *Cancer Res*. 2007;67:876-880.
10. Environmental Protection Agency. The Toxics Release Inventory (TRI) and Factors to Consider When Using TRI Data. Available at: http://www.epa.gov/tri/triprogram/FactorsToConPDF.pdf. Accessed March 22, 2012.
11. US Census Bureau. Introduction to Census 2000 Data Products. Available at: http://www.census.gov/prod/2001pubs/mso-01icdp.pdf. Accessed March 22, 2012.
12. Flowers C, Fedewa S, Chen A, et al. Disparities in the early adoption of chemoimmunotherapy for diffuse large B-cell lymphoma in the United States. *Cancer Epidemiol Biomarkers Prev*. 2012;21:1520-1430.
13. Abouyabis AN, Shenoy PJ, Lechowicz MJ, Flowers CR. Incidence and outcomes of the peripheral T-cell lymphoma subtypes in the United States. *Leuk Lymphoma*. 2008;49:2099-2107.
14. Shenoy P, Maggioncalda A, Malik N, Flowers CR. Incidence patterns and outcomes for Hodgkin lymphoma patients in the United States [serial online]. *Adv Hematol* 2011;2011:725219.
15. Shenoy P, Malik N, Sinha R, et al. Racial differences in the presentation and outcomes of chronic lymphocytic leukemia and variants in the United States. *Clin Lymphoma Myeloma Leuk*. 2011;11:498-506.
16. Shenoy PJ, Malik N, Nooka A, et al. Racial differences in the presentation and outcomes of diffuse large B-cell lymphoma in the United States [published online ahead of print December 22, 2010]. *Cancer*. 2010.
17. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Incidence—SEER 13 Registries Research Data, November 2010 Submission (1973-2008) <Katrina/Rita Population Adjustment> Linked to County Attributes—Total US Bethesda, MD: National Cancer Institute; 2009.
18. Morton LM, Turner JJ, Cerhan JR, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). *Blood*. 2007;110:695-708.
19. Environmental Systems Research Institute [ESRI]. Common questions. Available at: http://www.esri.com/software/coder/common-questions. Accessed January 9, 2012.
20. US Census Bureau. Census 2000 TIGER/Line Files: TIGER/Line Shapefiles and TIGER/Line Files. Available at: http://www.census.gov/geo/www/tiger/tiger2k/tgr2000.html. Accessed January 9, 2012.
21. Waller L, Gotway C. Applied Spatial Statistics for Public Health Data. New York: John Wiley & Sons, Inc.; 2004.
22. Kleinbaum D, Kupper L, Nizam A, Muller K. Applied Regression Analysis and Other Multivariable Methods. Pacific Grove, CA: Thomson Brooks/Cole; 2008.
23. Chang CM, Wang SS, Dave BJ, et al. Risk factors for non-Hodgkin lymphoma subtypes defined by histology and t(14;18) in a population-based case-control study. *Int J Cancer*. 2011;129:938-947.
24. Ambinder AJ, Shenoy PJ, Malik N, Maggioncalda A, Nastoupil LJ, Flowers CR. Exploring risk factors for follicular lymphoma [serial online]. *Adv Hematol* 2012:626035, 2012.
25. Wallace LA. Major sources of benzene exposure. *Environ Health Perspect*. 1989;82:165-169.
26. Morton LM, Hartge P, Holford TR, et al. Cigarette smoking and risk of non-Hodgkin lymphoma: a pooled analysis from the International Lymphoma Epidemiology Consortium (InterLymph). *Cancer Epidemiol Biomarkers Prev*. 2005;14:925-933.
27. Brauner EV, Sorensen M, Gaudreau E, et al. A prospective study of organochlorines in adipose tissue and risk of non-Hodgkin lymphoma. *Environ Health Perspect*. 2012;120:105-111.
28. The Environmental Integrity Project and the Galeston-Houstohn Association for Smog Prevention. Who's Counting? The Systematic Underreporting of Toxic Air Emissions. Washington, DC: US Environmental Protection Agency; 2004.
29. Smith MT, Jones RM, Smith AH. Benzene exposure and risk of non-Hodgkin lymphoma. *Cancer Epidemiol Biomarkers Prev*. 2007;16:385-391.
30. Cocco P, t'Mannetje A, Fadda D, et al. Occupational exposure to solvents and risk of lymphoma subtypes: results from the Epilymph case-control study. *Occup Environ Med*. 2010;67:341-347.
31. Vlaanderen J, Lan Q, Kromhout H, Rothman N, Vermeulen R. Occupational benzene exposure and the risk of lymphoma subtypes: a meta-analysis of cohort studies incorporating 3 study quality dimensions. *Environ Health Perspect*. 2011;119:159-167.
32. Kane EV, Newton R. Benzene and the risk of non-Hodgkin lymphoma: a review and meta-analysis of the literature. *Cancer Epidemiol*. 2010;34:7-12.

33. Linos A, Blair A, Gibson RW, et al. Leukemia and non-Hodgkin's lymphoma and residential proximity to industrial plants. *Arch Environ Health*. 1991;46:70-77.

34. Ramis R, Vidal E, Garcia-Perez J, et al. Study of non-Hodgkin's lymphoma mortality associated with industrial pollution in Spain, using Poisson models [serial online]. *BMC Public Health*. 2009;9:26.

35. Johnson KC, Pan S, Fry R, Mao Y. Residential proximity to industrial plants and non-Hodgkin lymphoma. *Epidemiology*. 2003;14:687-693.

36. Ramis R, Diggle P, Boldo E, Garcia-Perez J, Fernandez-Navarro P, Lopez-Abente G. Analysis of matched geographical areas to study potential links between environmental exposure to oil refineries and non-Hodgkin lymphoma mortality in Spain [serial online]. *Int J Health Geogr*. 2012;11:4.

37. Viel JF, Floret N, Deconinck E, Focant JF, De Pauw E, Cahn JY. Increased risk of non-Hodgkin lymphoma and serum organochlorine concentrations among neighbors of a municipal solid waste incinerator. *Environ Int*. 2011;37:449-453.

38. Wheeler DC, Ward MH, Waller LA. Spatial-temporal analysis of cancer risk in epidemiologic studies with residential histories [serial online]. *Ann Assoc Am Geogr*. 2012;102:1049-1057.

39. Dreiher J, Novack V, Barachana M, Yerushalmi R, Lugassy G, Shpilberg O. Non-Hodgkin's lymphoma and residential proximity to toxic industrial waste in southern Israel. *Haematologica*. 2005;90:1709-1710.

40. Dahlgren J, Klein J, Takhar H. Cluster of Hodgkin's lymphoma in residents near a non-operational petroleum refinery. *Toxicol Ind Health*. 2008;24:683-692.