ARTICLE COVERSHEET

LWW_CONDENSED(8.125X10.875)

SERVER-BASED

Article : SMJ50273
Creator : jtigas
Date : Wednesday August 24th 2016
Time : 10:50:30
Number of Pages (including this page) : 11

Relations Between Residential Proximity to EPA-Designated Toxic Release Sites and Diffuse Large B-Cell Lymphoma Incidence

AQ1 Catherine Bulka, MPH, Loretta J. Nastoupil, MD, Jean L. Koff, MD, Leon Bernal-Mizrachi, MD, Kevin Ward, PhD, Jessica N. Williams, BS, A. Rana Bayakly, MPH, Jeffrey M. Switchenko, PhD, Lance A. Waller, PhD, and Christopher R. Flowers, MD, MS

Objectives: Examining the spatial patterns of diffuse large B-cell lymphoma (DLBCL) incidence and residential proximity to toxic release locations may provide insight regarding environmental and sociodemographic risk factors.

Methods: We linked and geocoded cancer incidence data for the period 1999–2008 from the Georgia Comprehensive Cancer Registry with population data from the US Census and the Environmental Protection Agency's Toxics Release Inventory. We conducted cluster analyses and constructed Poisson regression models to assess DLBCL incidence as a function of mean distance to the toxic release sites.

Results: In total, 3851 incident DLBCL cases occurred among adults residing in Georgia between 1999 and 2008. Significant focal clustering was observed around 57% of ethylene oxide sites, 5% of benzene sites, 9% of tetrachloroethylene sites, 7% of styrene sites, 10% of formaldehyde sites, 5% of trichloroethylene sites, and 10% of all release sites. Mean distance to sites was significantly associated with DLBCL risk for all chemicals.

Conclusions: Proximity to Toxics Release Inventory sites can be linked to increased DLBCL risk as assessed through focal clustering and Poisson regression, and confirmatory studies using geospatial mapping can aid in further specifying risk factors for DLBCL.

Key Words: diffuse large B-cell lymphoma, epidemiology, lymphoma, non-Hodgkin lymphoma

From the Division of Epidemiology and Biostatistics, University of Illinois at Chicago School of Public Health, Chicago, the University of Texas MD Anderson Cancer Center, Houston, the Departments of Hematology and Oncology and Biostatistics and Bioinformatics Shared Resource, Winship Cancer Institute, and the Departments of Epidemiology and Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, and the Georgia Department of Public Health, Atlanta.

Correspondence to Dr Christopher Flowers, Department of Hematology and Oncology/Winship Cancer Institute, Emory University, 1365 Clifton Rd, NE, Building B, Suite 4302, Atlanta, GA 30322. E-mail: crflowe@emory.edu.

The work described herein was supported in part by National Cancer Institute grant no. R21 CA158686 and an American Society of Hematology Clinical Scholars Award.

AQ2 L.J.N. has received honoraria from Celgene and Genentech. L.B.M. has patents planned, pending, or issued with compensation to him from Empire Genomics; he has served on the speakers' bureau of Celgene; and he has consulted to X. L.A.W. has received grants from the *National Institute of Environmental Health Sciences, the National Center for Advancing Translational Sciences, the* National Heart, Lung, and Blood Institute, the US Environmental Protection Agency, *and* Ecological Associates Inc; and he has served as an expert witness for Pope McGlamery. C.R.F. has received a grant from the National Cancer Institute (R21 CA158686) and grants from AbbVie, Acerta, Celgene, Genentech, Gilead Sciences, Infinity Pharmaceuticals, Janssen, Millennium/Takeda, Onyx Pharmaceuticals, Pharmacyclics, Spectrum and TG Therapeutics; he has received an American Society of Hematology Clinical Scholars Award; he has received consulting fees from Seattle Genetics and OptumRx and is an unpaid consultant to Genentech and Celgene. The remaining authors have no financial relationships to disclose and no conflicts of interest to report.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (http://sma.org/smj-home)

Accepted May 3, 2016.

0038-4348/0–2000/109-00

DOI: 10.14423/SMJ.0000000000000545

Diffuse large B-cell lymphoma (DLBCL) is the most common non-Hodgkin lymphoma (NHL) in the Western world, comprising approximately one-third of adult lymphomas. We previously examined spatial patterns of NHL incidence, the association between NHL incidence and distance to benzene release sites,[1] and racial differences in incidence and outcomes for DLBCL and other NHL subtypes.[1–3] Biologically, exposure to industrial volatile organic compounds such as benzene produces chromosomal aberrations and genetic changes. Some effects can occur at air levels of ≤1 ppm, suggesting that even low levels of chronic exposure can be harmful.[4] Our group

Key Points

- We found significant focal clustering of diffuse large B-cell lymphoma (DLBCL) risk around Georgia toxic sites that release chemicals such as ethylene oxide, benzene, and tetrachloroethylene.
- Based on Poisson regression results, we discovered significant associations between mean distance to sites and DLBCL risk.
- Our results can assist in identifying candidate organic compounds to measure directly in epidemiological studies of individuals at risk for DLBCL.

was intrigued by reports that DLBCL risk is associated with residential proximity to industrial facilities,[5] data indicating that race and socioeconomic status are strongly associated with residential proximity to Toxics Release Inventory (TRI) facilities,[6–9] and our findings of racial differences in DLBCL incidence, presentation, and outcome.[2,10,11] As such, we sought to examine socioeconomic and racial differences in the spatial epidemiology of DLBCL and assess the relation between DLBCL incidence and residential distance from TRI exposure sites.

Methods

We linked geocoded data on toxic release sites in Georgia between 1988 and 1998 from the Environmental Protection Agency's (EPA's) TRI, DLBCL incidence from the Georgia Comprehensive Cancer Registry (GCCR) from 1999 to 2008, and census tract population and demographic data from the 2000 US Census. This study was approved by the Emory University and the Georgia Department of Public Health institutional review boards.

Toxic Release Site Data

The EPA requires that facilities meeting thresholds defined by Section 313 of the Emergency Planning and Community Right-to-Know Act (42 USC 11001–11050) report annually their disposal or releases for listed toxic chemicals.[12] Launched in 1987, the TRI captures release information about certain chemicals, including quantities, media type, and geographic coordinates of releases from sources such as manufacturing facilities, service businesses, and federal facilities.[12] We measured and geocoded data for TRI sites in Georgia.

Census Tract Population and Demographic Data

During the 2000 US Census, there were 1618 census tracts within the state of Georgia. Population and demographic data were available from US Census Summary File 1 for 1616 of these tracts.[13] We used the percentage of individuals in each census tract living below the federal poverty level as a metric for socioeconomic status (SES) and determined the racial composition of each tract. We used census data on median year moved into residence (MYMI) derived from a sample of individuals residing within a census tract to estimate the length of time that residents were living in their current home as a surrogate for duration of residential exposure to TRI sites.

GCCR Data

GCCR provided data for all incident DLBCL cases among adults (20 years old and older) residing in Georgia at diagnosis during the period 1999–2008 using *International Classification of Diseases for Oncology, First Revision* codes from the World Health Organization–based classification of lymphoid neoplasms for epidemiologic research from InterLymph.[14] We previously examined national population–level incidence data for DLBCL and other NHL subtypes.[2,3,10,11,15] SEER*Stat (Surveillance, Epidemiology, and End Results program) version 7.05 (National Cancer Institute, Bethesda, MD) was used to access the SEER 13 Registries Database and standardize the DLBCL incidence rates from Georgia to national DLBCL incidence rates.[16] Age-, sex-, and race-specific crude DLBCL incidence rates were obtained for the 1999–2008 period to standardize Georgia census tract incidence data by age, sex, and race.

Subjects in the GCCR without age, sex, or race information were excluded from all of the analyses. We used four major race categories (white, African American, American Indian/Alaskan Native, and Asian or Pacific Islander); we excluded subjects whose race was categorized as "other" or "unknown." Successfully geocoded cases were aggregated to the census tract level.[17] Standardized incidence ratios (SIRs) were calculated for each census tract by dividing the number of observed cases by expected cases.

Geographic Data and Spatial Analyses

We used the geographic information system (GIS) ArcGIS 10 (ESRI, Redlands, CA) to examine the spatial distribution of TRI sites and SIRs by census tract. Census tract shapefiles were obtained from the US Census Bureau's 2000 TIGER/Line files.[13] GIS software calculated distances from the 1618 census tract centroids to each toxic release site. Choropleth maps depict the DLBCL SIRs by census tract. Locations of TRI sites were overlaid on the SIR maps. Choropleth maps also were created to depict DLBCL incidence, percentage of individuals living below the federal poverty level, and racial composition by census tract in relation to TRI sites. Because sparse data can pose statistical problems in small area analyses, spatial smoothing also was performed to allow census tracts to borrow information from spatial neighbors and produce more stable estimates of SIRs while maintaining geographic precision. Specifically, spatial Empirical Bayes smoothing was performed on the SIR values using GeoDa 1.01 (Luc Anselin, Tempe, AZ).[18] We defined "neighbors" as census tracts with common borders or vertices and constructed choropleth maps of SIRs for DLBCL.

To assess the spatial correlation of SIRs, we conducted global and local spatial analyses. A global measure of spatial autocorrelation, Moran's I, was calculated for SIR patterns of DLCBL using GeoDa. To measure spatial autocorrelation at the local scale, local Moran's I, also called local indicators of spatial autocorrelation (LISA),[18] was calculated for SIR patterns of DLBCL using GeoDa. LISA cluster maps were created to identify the locations of any significant clusters of SIRs. Both the global and local spatial statistics based significance on 999 Monte Carlo simulations, with neighbors defined as above.[18]

We tested for focal clustering, using the observed cases and the population at risk in each census tract to test the null hypothesis that there was no spatial clustering near the TRI sites (the foci). The Lawson-Waller score test was used to assess individually each of the sites for focal clustering of DLBCL.[18] Each census tract was scored for the difference between the observed

and expected counts of DLBCL, weighted by inverse distance from each census tract centroid to the focus.[18] The Lawson-Waller test statistic for site i is the following:

$$T_i = \sum_{j=1}^{n} W_{ij}\left(c_j - n_j \frac{C}{n}\right)$$

where c_j is the number of observed cases in census tract j, n_j is the population of census tract j, W_{ij} is the inverse distance between census tract j and site i, C is the total number of cases, and n is the total population at risk. The standard normal distribution was used to estimate upper-tail P values. Because this test was conducted once for each release site, we adjusted our significance level using the Bonferroni correction.[19]

Poisson regression models were constructed under the assumption that the number of observed incident cases for each census tract had a Poisson distribution where the expected number of cases for that census tract was based on local age, sex, and race demographics, and the explanatory variable was mean distance from the TRI sites. The census tract measures of SES and MYMI were assessed as potential confounders and/or effect modifiers. The models were estimated with the maximum likelihood method using SAS version 9.3 (SAS Institute, Cary, NC).

Results

From 1988 to 1998, Georgia facilities reported the release of 1,3 butadiene (3 sites), 2,4-D (1 site), benzene (19 sites), ethylene oxide (7 sites), formaldehyde (60 sites), pentachlorophenol (5 sites), styrene (86 sites), tetrachloroethylene (33 sites), and trichloroethylene (TCE, 40 sites). Total releases per site, calculated as the sum of fugitive air releases, stack air releases, and surface water discharges were 250 lb of 2,4-D and ranged from 52 to 3,830,097 lb of benzene; 171,437 to 216,659 lb of 1,3 butadiene; 4220 to 581,077 lb of ethylene oxide; 5 to 872,835 lb of formaldehyde; 164 to 3845 lb of pentachlorophenol; 2 to 4,472,334 lb of styrene; 5 to 1,575,644 lb of tetrachloroethylene; and 5 to 3,730,069 lb of TCE. Of the 4316 incident cases of DLBCL in our original dataset, 3857 cases were successfully geocoded to census tracts. Among these, 3581 (92.8%) cases had age, sex, and race information available to be classified into the US Census race categories (Table 1). We mapped SIRs for DLBCL with location data for each TRI site (Fig. 1). Two census tracts were not included because of a lack of demographic data.

Table 1. Study population demographics

	Cases (n = 3851)		Georgia population (N = 5,582,707)[a]	
Variable	**N**	**%**	**N**	**%**
Age group, y				
20–59	1560	40.5	4,521,891	81.0
≥ 60	2291	59.5	1,060,816	19.0
Sex				
Female	1779	46.2	2,907,559	52.1
Male	2072	53.8	2,675,148	47.9
Race				
White	3304	78.8	3,922,068	70.3
African American	757	19.7	1,521,316	27.22
American Indian/Alaskan Native	—[b]	0.0	15,326	0.3
Asian or Pacific Islander	59	1.5	123,997	2.2
MYMI			1989	

MYMI, median year moved into residence.

[a]Based on 1616 census tracts in Georgia from the 2000 US Census.

[b]Data suppressed.

Spatial Analyses

We found no evidence of global spatial correlation of SIRs among all of the cases or cases restricted to specific age, sex or race subgroups at the $\alpha = 0.0071$ level (Bonferroni adjusted for seven comparisons); however, LISA cluster maps of SIRs (Fig. 2) demonstrated some local 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. Clustering of high SIRs occurred in the metropolitan Atlanta area for all DLBCL cases and for each population subgroup (Fig. 2). High-high clustering of DLBCL was primarily located in the northern half of the state and low-low clustering in the southern half.

Among all of the cases, mean distance to formaldehyde release sites had the strongest effect, with a 0.58% decrease in DLBCL risk expected for every mile that the distance to formaldehyde release sites increased ($P < 0.0001$). Tables 2 to 5 show the results of the Poisson regression analyses. The relative risk is equal to $\exp(\beta_1)$; in other words, increasing mean distance to TRI sites by 1 mi decreases the risk of DLBCL by $1 - \exp(\beta_1)$. Mean distance from 1,3-butadiene, 2,4-D, benzene, ethylene oxide, formaldehyde, styrene, tetrachloroethylene, and TCE release sites associated significantly with DLBCL incidence, whereas mean distance from pentachlorophenol sites did not. No statistically significant Poisson coefficients emerged for the interaction term between mean distance to TRI sites and the percentage of individuals living below the poverty level or MYMI; as such, these variables were removed from multivariable models. In addition, comparing models that included these variables with those that did not suggested that these were not confounders in the relation between mean distance to TRI sites and DLBCL incidence. Notably, the effect of mean distance from all of the release sites on DLBCL incidence was strongest for African Americans (Table 3). Relations among percentage of individuals living below the poverty level, racial composition of census tracts, DLBCL incidence, and proximity to TRI sites are displayed using choropleth maps in Figures 3 and 4.

Using the Lawson-Waller score test, all nine toxic release exposures showed some evidence of focal clustering of DLBCL

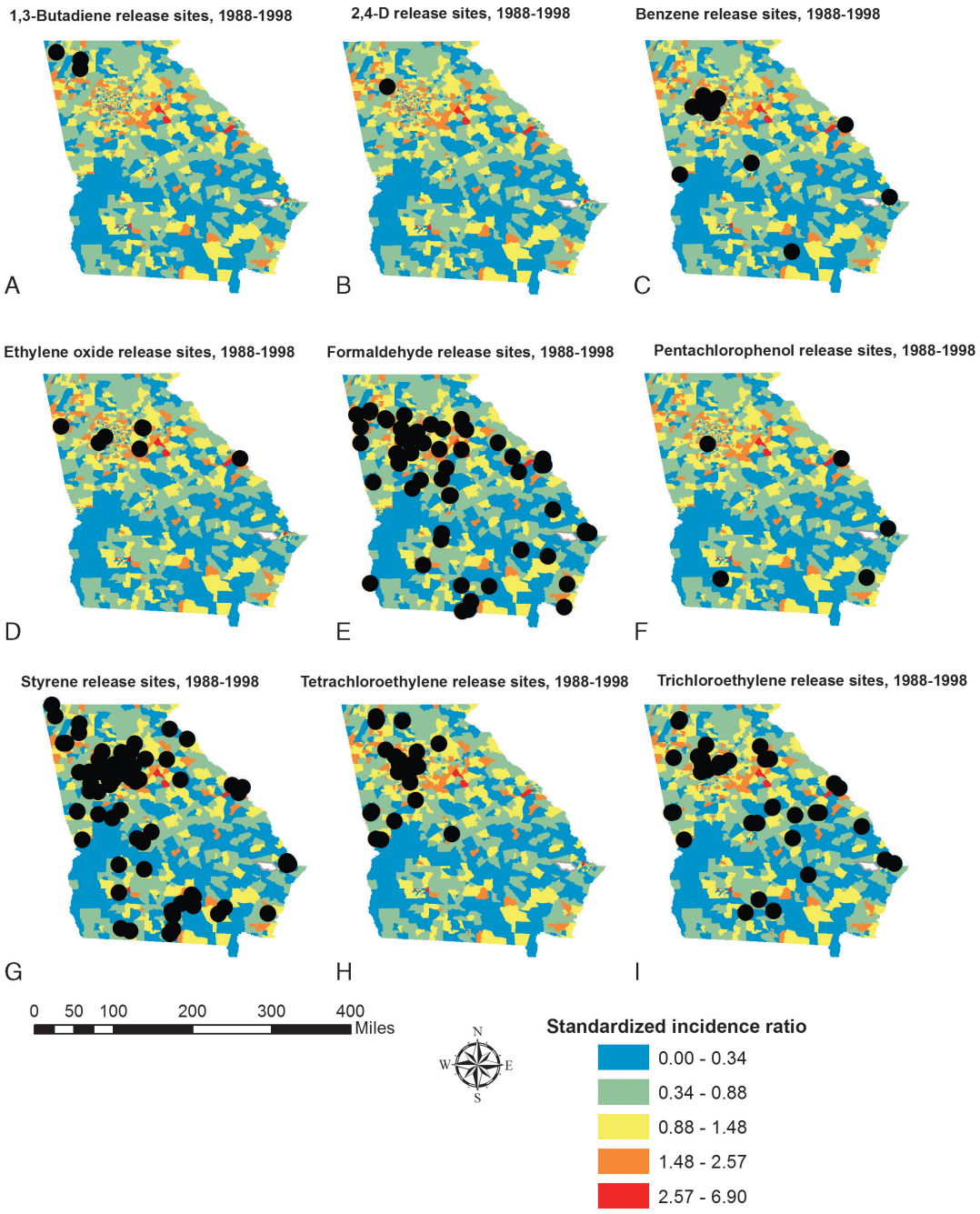

Fig 1 4/C

Fig. 1. Standardized incidence ratio of diffuse large B-cell lymphoma by census tract with Environmental Protection Agency–designated toxic release sites.

T6 (Supplementary Table S1 http://links.lww.com/SMJ/xxx, Table 6). Significant focal clustering was observed around TRI sites, and the rates of clustering around sites differed by race.

Discussion

Numerous studies support associations between occupational toxic exposures and NHL incidence,[20–22] although considerable controversy remains regarding residential exposures.[5,23,24] Our study examined relations between incidence of DLBCL at the census tract level and toxic release sites to determine whether DLBCL incidence varies with distance from these sites. We identified areas of spatial clustering of DLBCL, which appeared to be high in the metropolitan Atlanta region and low in the southern areas of the state. All nine toxic release exposures showed some evidence for focal clustering, but not every site harbored a cluster of increased DLBCL incidence. The potential reasons for significantly increased DLBCL incidence around certain sites but not others include spurious associations,

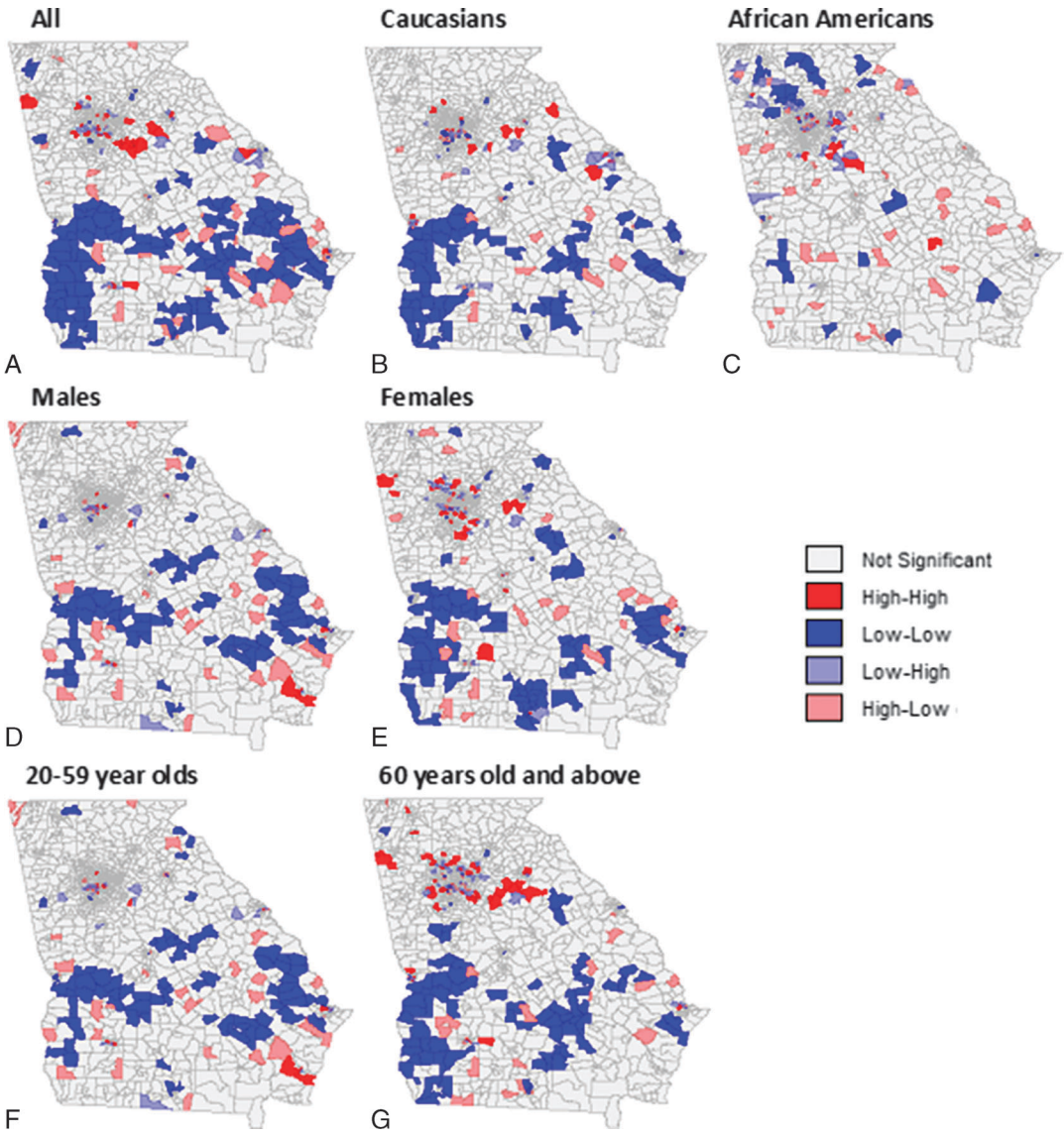

Fig 2 4/C

Fig. 2. **Local indicators of spatial autocorrelation clustering of diffuse large B-cell lymphoma incidence by census tract.**

Table 2. Poisson regression results

	All		
Exposure	**β_1**	**exp(β_1)**	***P***
1,3-Butadiene	−0.0012	0.9988	<0.001
2,4-D	−0.0021	0.9979	<0.001
Benzene	−0.0032	0.9968	<0.001
Ethylene oxide	−0.0032	0.9968	<0.001
Formaldehyde	−0.0058	0.9942	<0.001
Pentachlorophenol	0.0017	1.0017	0.002
Styrene	−0.0054	0.9946	<0.001
Tetrachloroethylene	−0.0027	0.9973	<0.001
Trichloroethylene	−0.0048	0.9952	<0.001

Table 3. Poisson regression results for African Americans and whites

	African Americans			Whites		
Exposure	**β_1**	**exp(β_1)**	***P***	**β_1**	**exp(β_1)**	***P***
1,3-Butadiene	−0.0028	0.9972	<0.001	−0.0016	0.9984	<0.001
2,4-D	−0.0031	0.9969	<0.001	−0.0019	0.9981	<0.001
Benzene	−0.0049	0.9951	<0.001	−0.0033	0.9967	<0.001
Ethylene oxide	−0.0044	0.9956	<0.001	−0.0029	0.9971	<0.001
Formaldehyde	−0.0091	0.9909	<0.001	−0.0051	0.9949	<0.001
Pentachlorophenol	0.0037	1.0037	0.021	0.0015	1.0015	0.010
Styrene	−0.0075	0.9925	<0.001	−0.0049	0.9951	<0.001
Tetrachloroethylene	−0.0037	0.9963	<0.001	−0.0025	0.9975	<0.001
Trichloroethylene	−0.0067	0.9933	<0.001	−0.0044	0.9956	<0.001

Table 4. Poisson regression results for women and men

	Women			Men		
Exposure	β_1	exp(β_1)	***P***	β_1	exp(β_1)	***P***
1,3-Butadiene	−0.0016	0.9984	<0.001	−0.0020	0.9980	<0.001
2,4-D	−0.0019	0.9981	<0.001	−0.0023	0.9977	<0.001
Benzene	−0.0033	0.9967	<0.001	−0.0039	0.9961	<0.001
Ethylene oxide	−0.0028	0.9972	<0.001	−0.0035	0.9965	<0.001
Formaldehyde	−0.0057	0.9943	<0.001	−0.0058	0.9942	<0.001
Pentachlorophenol	0.0012	1.0012	0.135	0.0022	1.0022	0.004
Styrene	−0.0051	0.9949	<0.001	−0.0056	0.9944	<0.001
Tetrachloroethylene	−0.0025	0.9975	<0.001	−0.0029	0.9971	<0.001
Trichloroethylene	−0.0045	0.9955	<0.001	−0.0051	0.9949	<0.001

differences in levels of exposure, and differences in other covariates influencing risk. We have attempted to address the latter by adjusting for age, sex, and race in risk estimates and examining SES and MYMI as potential confounders, but other known and unknown unadjusted factors may influence DLBCL risk. Notably, release sites with significant focal clustering tended to occur in areas with higher population density, primarily in the metropolitan Atlanta region, even though our analyses adjusted for this factor.

We based our study on the underlying assumption that increased exposure to the chemicals emerging from TRI facilities can produce an elevated risk of lymphoma and used mean distance between residence and emission facilities as an indicator of exposure. All of the chemicals investigated are volatile organic compounds for which indoor sources and outdoor sources may contribute to personal exposure. Outdoor sources include point sources (eg, TRI facilities), nonpoint, on-road, secondary, and background sources. In addition, chemical exposures from TRI facilities depend not only on distance but also on wind speed, direction, water, and other sources. Accordingly, these data should not be used alone to determine environmental policy regarding DLBCL. Instead, our results aid in identifying candidate organic compounds that can be measured directly in epidemiological studies of individuals at risk and can aid in identifying populations in which such measurement should be performed.

One key strength of our study is the use of publicly available data from the EPA and the US Census Bureau. The racial and SES composition of the state of Georgia, which varied by region, provided a distinct opportunity to investigate relations among race, TRI exposure, and DLBCL incidence. Our study, however, was limited by the use of aggregated data, and the associations observed at the census tract level 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. Additional studies are needed that directly measure individuals' blood toxin levels and examine their association with DLBCL incidence. Such studies could be targeted toward populations designated to be at increased risk by residential proximity and could provide corroborating evidence for the relation between these exposures and DLBCL. For approximately 18% of patients with DLBCL, accurate location data were missing, thereby precluding their inclusion in the geospatial analyses. We compared the DLBCL cohort included in the geospatial analysis with patients with DLBCL from the original full dataset of the GCCR and found no significant differences in race ($P = 0.87$) or sex ($P = 0.43$) using χ^2 tests, or age ($P = 0.13$) using a *t* test comparing those who were geocoded with a census tract and those who were not geocoded.

Another potential weakness of our study was the lack of quantitative exposures and temporal analyses. We attempted to control for time and latency by including only toxic release sites from 1988 to 1998 and using incidence and population data from 1999 to 2008, as well as assessing MYMI as a potential confounder and effect modifier. In addition, our estimate of exposure, mean distance to TRI sites for each census tract, does not include the quantity of chemicals released, which could improve estimates of cumulative exposure. In addition, mean distance may yield a lower exposure for tracts close to one site but far from other sites, which may result in an unreasonably low estimate for exposure. Future directions should include development of methods for estimating cumulative exposure as a function of distance to site, quantity of release at each site, and weather factors such as average wind patterns and speeds. Other possible sources of bias include changes in the population of Georgia from 1999 to 2008 and the use of 2000 US Census data as denominators for incidence rates. Because of these and other limitations, this work cannot be used to specifically delineate exposures that cause DLBCL.

Although causative etiologies remain unknown for most lymphomas, factors that influence the risk of developing lymphoma include genetics; environmental factors such as exposures to ultraviolet radiation, pesticides, or hair dyes; and comorbid diseases or their treatments, especially those that cause significant immunosuppression.[25] Advances in genomics

Table 5. Poisson regression results for 20- to 59-year-olds and ≥60-year-olds

	20–59 y old			≥60 y old		
Exposure	β_1	exp(β_1)	***P***	β_1	exp(β_1)	***P***
1,3-Butadiene	−0.0020	0.9980	<0.001	−0.0017	0.9983	<0.001
2,4-D	−0.0023	0.9977	<0.001	−0.0020	0.9980	<0.001
Benzene	−0.0041	0.9959	<0.001	−0.0034	0.9966	<0.001
Ethylene oxide	−0.0036	0.9964	<0.001	−0.0030	0.9970	<0.001
Formaldehyde	−0.0065	0.9935	<0.001	−0.0054	0.9946	<0.001
Pentachlorophenol	0.0021	1.0021	0.027	0.0015	1.0015	0.025
Styrene	−0.0060	0.9940	<0.001	−0.0051	0.9949	<0.001
Tetrachloroethylene	−0.0030	0.9970	<0.001	−0.0025	0.9975	<0.001
Trichloroethylene	−0.0054	0.9946	<0.001	−0.0045	0.9955	<0.001

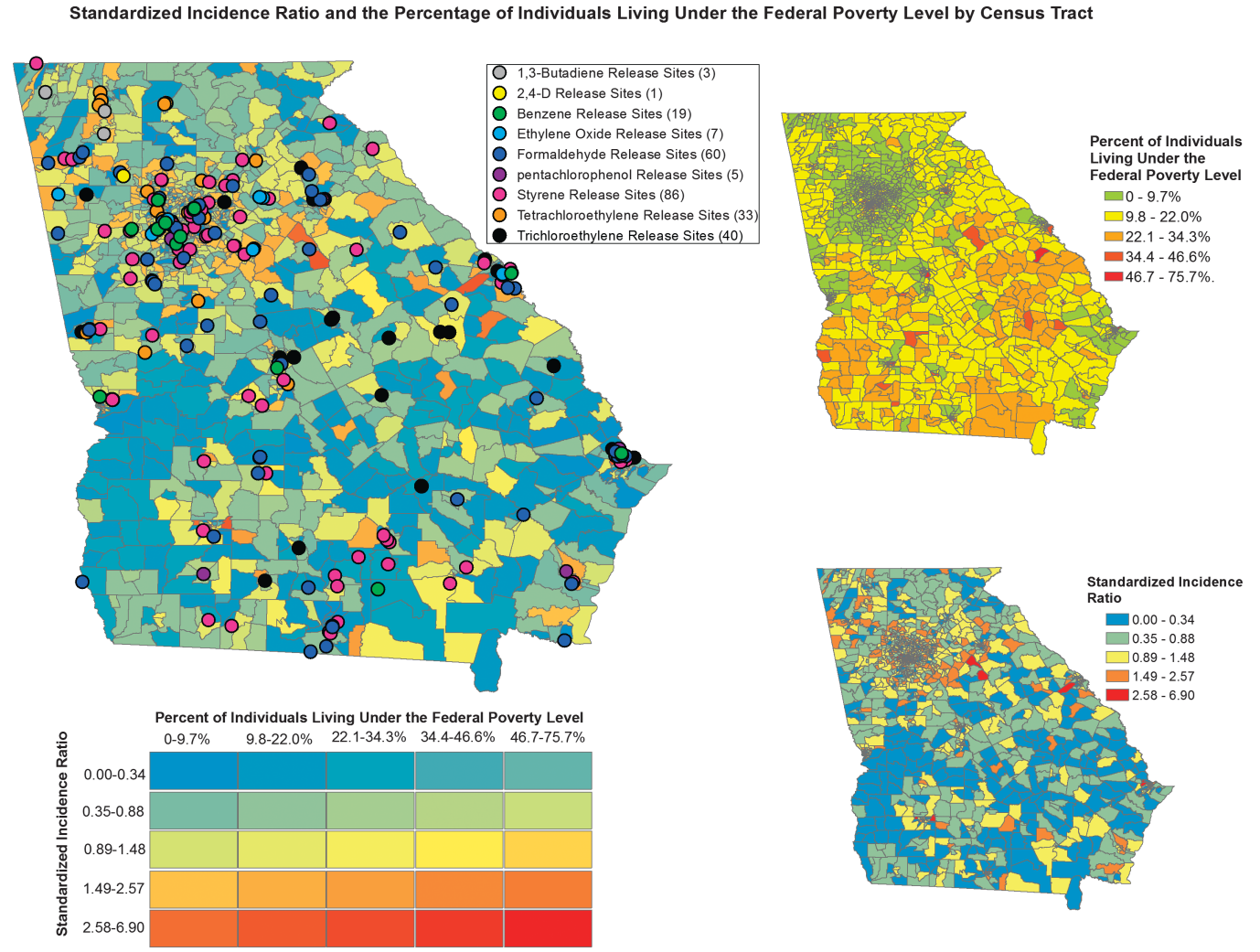

Fig 3 4/C

Fig. 3. Standardized incidence ratios of diffuse large B-cell lymphoma and percentage of individuals living below the federal poverty level by census tract.

support a genetic basis for etiologic commonality among lymphomas and for heterogeneity underlying distinct subtypes[26–30] and suggest that immune dysfunction is of greater etiologic importance for DLBCL than other lymphoma subtypes.[25] As a result, many studies of lymphomagenesis have focused on genes involved in inflammatory pathways and in lymphoid development.[31,32] Research also has shown that the relation between a toxic exposure and NHL risk may be modified by particular variants in immunoregulatory genes, suggesting a need for further studies of gene–environment interactions.[31]

Several case–control studies have explored environmental exposures potentially associated with DLBCL and NHL. Researchers obtained 20-year residential histories for 1321 NHL cases and 1057 controls in four US SEER centers and detected clusters of significantly elevated risk in one of the study areas (Los Angeles) at a lag time of 20 years ($P = 0.03$).[33] Another study examined the residential locations in four SEER regions during the 10 years before recruitment to characterize the impact that proximity to industrial facilities had on developing NHL. This study found that living within 2 mi of a lumber facility was associated with increased DLBCL risk, but it did not provide strong evidence that living near manufacturing industries increases NHL risk overall.[5] Although some studies have found associations between increased lymphoma risk and ultraviolet radiation, chemicals (particularly benzene and TCE), agricultural exposures, and hair dyes, debate remains regarding these relations generally and for DLBCL specifically.[34–40] Novel approaches are needed to integrate clinical and genomic data from lymphoma patients with residential exposure data to identify the effects of toxic exposures on molecular and cellular pathways. Early efforts in this arena are under way.[41]

The results of our study also lend insight into the possible interplay among race, environmental factors, and DLBCL. It has been well established that there are significant racial differences in DLBCL incidence. Based on data from 37,009 cases of DLBCL in the SEER registry, whites in the United States have an age-adjusted incidence of 7.36/100,000 in the population and African Americans have an incidence of 4.88/100,000, with African

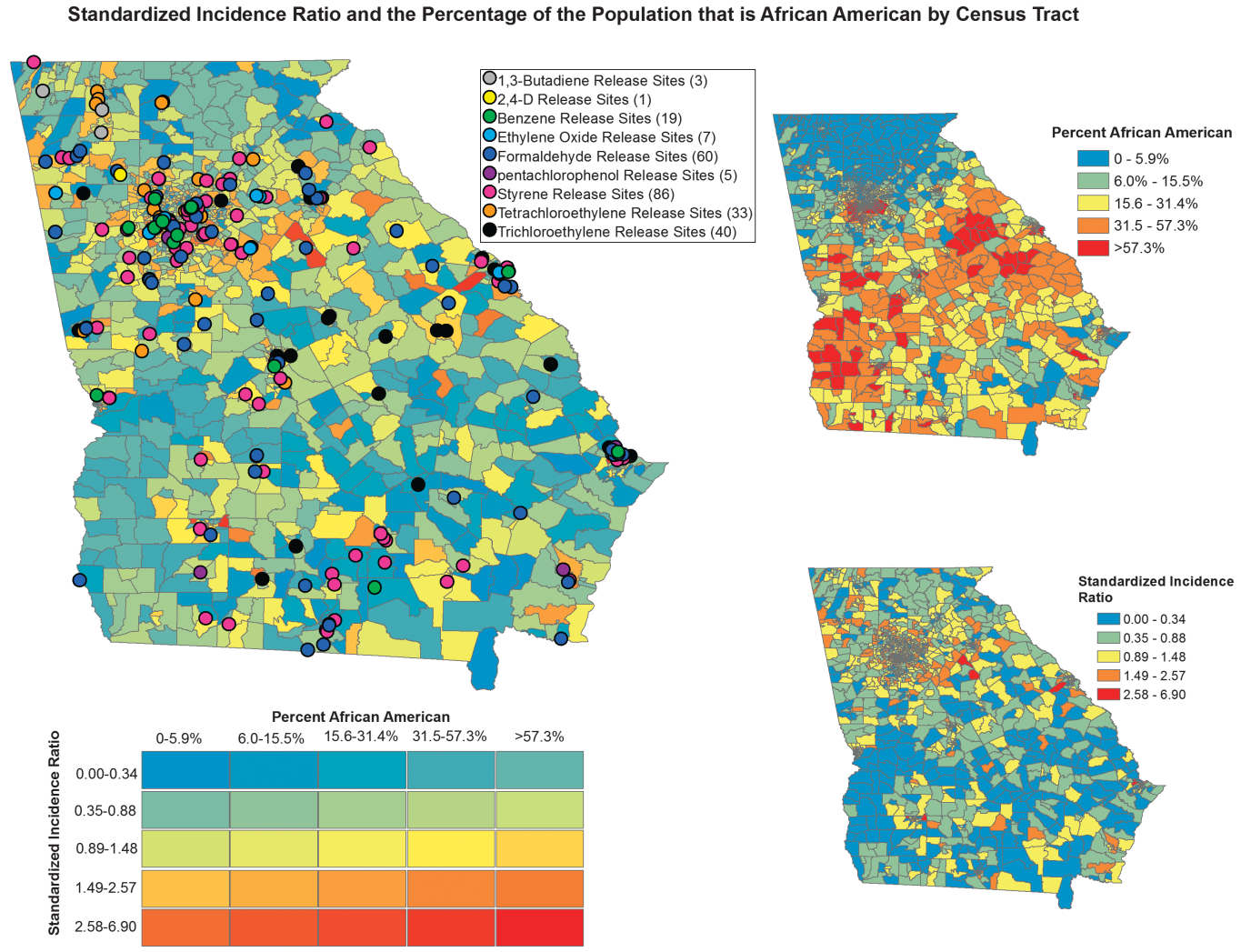

Fig 4 4/C

Fig. 4. Standardized incidence ratios of diffuse large B-cell lymphoma and percentage of the African American population by census tract.

American patients more commonly presenting with DLBCL at 60 years old and younger (65% vs 37%); this indicates that there are racial differences in the rates and patterns of DLBCL development.[2] In addition, racial and socioeconomic disparities have been described with regard to exposure to airborne toxins, with benzene being one of the largest contributors.[6] The magnitude of disparity is highest in the poorest and most highly concentrated areas in which African Americans reside.

Although the effect of mean distance from certain TRIs on DLBCL incidence was stronger for African Americans and the patterns of focal clustering of DLBCL cases around EPA TRI sites differed by race, these patterns did not appear to be related to differences in racial population distributions across Georgia. These data suggest racial differences in the magnitude of exposure, genetic susceptibility to toxic exposures, or other factors that can contribute to increased risk. This study is the first to our knowledge to assess the relation among multiple toxic release exposures, race, and the incidence of a specific NHL subtype and can provide a foundation for future work.

Table 6. Number of significant focal clusters among total number of TRI sites per chemical and by racial group

Chemical	Significant clustering (Overall %)	Significant clustering (white %)	Significant clustering (African American %)
1,3-Butadiene	3/3 (100)	0/3 (0)	2/3 (67)
2,4-D	1/1 (100)	0/1 (0)	0/1 (0)
Benzene	1/19 (5)	4/19 (21)	5/19 (26)
Ethylene oxide	4/7 (57)	4/7 (57)	6/7 (86)
Formaldehyde	6/60 (10)	2/60 (3)	3/60 (5)
Pentachlorophenol	0/5 (0)	0/5 (0)	1/5 (20)
Styrene	6/86 (7)	7/86 (8)	14/86 (16)
Tetrachloroethylene	3/33 (9)	3/33 (9)	8/33 (24)
Trichloroethylene	2/40 (5)	1/40 (3)	5/40 (13)

TRI, Toxics Release Inventory.

References

1. Bulka C, Nastoupil LJ, McClellan W, et al. Residence proximity to benzene release sites is associated with increased incidence of non-Hodgkin lymphoma. *Cancer* 2013;119:3309–3317.
2. 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. *Cancer* 2011;117:2530–2540.
3. Shenoy PJ, 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.
4. 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.
5. 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.
6. James W, Jia C, Kedia S. Uneven magnitude of disparities in cancer risks from air toxics. *Int J Environ Res Public Health* 2012;9:4365–4385.
7. Perlin SA, Wong D, Sexton K. Residential proximity to industrial sources of air pollution: interrelationships among race, poverty, and age. *J Air Waste Manag Assoc* 2001;51:406–421.
8. Morrens B, Bruckers L, Hond ED, et al. Social distribution of internal exposure to environmental pollution in Flemish adolescents. *Int J Hyg Environ Health* 2012;215:474–481.
9. Hun DE, Siegel JA, Morandi MT, et al. Cancer risk disparities between Hispanic and non-Hispanic white populations: the role of exposure to indoor air pollution. *Environ Health Perspect* 2009;117:1925–1931.
10. Flowers CR, Fedewa SA, Chen AY, et al. Disparities in the early adoption of chemoimmunotherapy for diffuse large B-cell lymphoma in the United States. *Cancer Epidemiol Biomarkers Prev* 2012;21:1520–1530.
11. Flowers CR, Shenoy PJ, Borate U, et al. Examining racial differences in diffuse large B-cell lymphoma presentation and survival. *Leuk Lymphoma* 2013;54:268–276.
12. Environmental Protection Agency. Factors to consider when using Toxics Release Inventory data. https://www.epa.gov/toxics-release-inventory-tri-program/factors-consider-when-using-toxics-release-inventory-data. Updated September 2, 2015. Accessed August 5, 2016.
13. US Census Bureau. Introduction to Census 2000 data products. http://www.census.gov/prod/2001pubs/mso-01icdp.pdf. Accessed July 28, 2016.
14. Abouyabis AN, Shenoy PJ, Lechowicz MJ, et al. Incidence and outcomes of the peripheral T-cell lymphoma subtypes in the United States. *Leuk Lymphoma* 2008;49:2099–2107.

AQ3

15. Surveillance Epidemiology and End Results (SEER) Program. SEER*Stat Database: Incidence—SEER 13 Registries Research Data, Nov 2010 Sub (1973–2008) < Katrina/Rita Population Adjustment > − Linked To County Attributes - Total U.S.
16. 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.
17. ESRI. Common questions. http://www.esri.com/training/main/common-questions. Accessed August 5, 2016.
18. Waller LA, Gotway C. *Applied Spatial Statistics for Public Health Data*. New York: Wiley, 2004.
19. Kleinbaum DG, Kupper LL, Nizam A, et al. *Applied Regression Analysis and Other Multivariable Methods*. Thomson Brooks/Cole: Independence, KY, 2008.
20. 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.
21. Rinsky RA, Hornung RW, Silver SR, et al. Benzene exposure and hematopoietic mortality: a long-term epidemiologic risk assessment. *Am J Ind Med* 2002;42:474–480.
22. Kipen HM, Cody RP, Crump KS, et al. Hematologic effects of benzene: a thirty-five year longitudinal study of rubber workers. *Toxicol Ind Health* 1988;4:411–430.
23. Ramis R, Diggle P, Boldo E, et al. Analysis of matched geographical areas to study potential links between environmental exposure to oil refineries and non-Hodgkin lymphoma mortality in Spain. *Int J Health Geogr* 2012;11:4.
24. Linos A, Blair A, Gibson RW, et al. Leukemia and non-Hodgkin's lymphoma and residential proximity to industrial plants. *Arch Environ Health* 1991;46:70–74.
25. Morton LM, Wang SS, Cozen W, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes. *Blood* 2008;112:5150–5160.
26. Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* 2011;43:830–837.
27. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A* 2012;109:3879–3884.
28. Love C, Sun Z, Jima D, et al. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet* 2012;44:1321–1325.
29. Morin RD, Mungall K, Pleasance E, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* 2013;122:1256–1265.
30. Zhang J, Grubor V, Love CL, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* 2013;110:1398–1403.
31. Colt JS, Rothman R, Severson RK, et al. Organochlorine exposure, immune gene variation, and risk of non-Hodgkin lymphoma. *Blood* 2009;113:1899–1905.
32. Rothman N, Skibola CF, Wang SS, et al. Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. *Lancet Oncol* 2006;7:27–38.
33. Wheeler DC, De Roos AJ, Cerhan JR, et al. Spatial-temporal analysis of non-Hodgkin lymphoma in the NCI-SEER NHL case–control study. *Environ Health* 2011;10:63.
34. Chang CM, Schroeder JC, Olshan AF, et al. A case–control study of tobacco use and other non-occupational risk factors for lymphoma subtypes defined by t(14; 18) translocations and bcl-2 expression. *Cancer Causes Control* 2010;21:1147–1154.
35. Fisher SG, Fisher RI. The epidemiology of non-Hodgkin's lymphoma. *Oncogene* 2004;23:6524–6534.
36. Friedberg JW, Fisher RI. Diffuse large B-cell lymphoma. *Hematol Oncol Clin North Am* 2008;22:941–952.
37. 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.
38. Lim U, Freedman DM, Hollis BW, et al. A prospective investigation of serum 25-hydroxyvitamin D and risk of lymphoid cancers. *Int J Cancer* 2009;124:979–986.
39. Zahm SH, Morton LM, Bernstein L, et al. Re: hair dye use, genetic variation in N-acetyltransferase 1 (NAT1) and 2 (NAT2), and risk of non-Hodgkin lymphoma, author response. *Carcinogenesis* 2008;29:1084–1085.
40. Zhang Y, Sanjose SD, Bracci PM, et al. Personal use of hair dye and the risk of certain subtypes of non-Hodgkin lymphoma. *Am J Epidemiol* 2008;167:1321–1331.
41. Zhang L, McHale CM, Rothman N, et al. Systems biology of human benzene exposure. *Chem Biol Interact* 2010;184:86–93.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

AQ1 = Please check if authors name are correctly captured for given names (in red) and surnames (in blue) for indexing after publication.

AQ2 = Please name the entity to which L.B.M. has consulted.

AQ3 = Ref 15 is incomplete/not able to be searched in present form. Pls provide URL or more concrete information so that readers can find the citation.

END OF AUTHOR QUERIES