Modeling historic environmental pollutant exposures and non-Hodgkin lymphoma risk

Joseph Boyle[1], **Mary H. Ward**[2], **James R. Cerhan**[3], **Nathaniel Rothman**[2], **David C. Wheeler**[1]

[1]Department of Biostatistics, Virginia Commonwealth University, Richmond, VA

[2]Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD

[3]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN

Abstract

Environmental exposures to chemicals are suspected risk factors for non-Hodgkin lymphoma (NHL), but few studies have assessed historic environmental risk factors. In this study, we estimated the associations between NHL and 1) historic environmental pollutant emissions from the Risk Screening Environmental Indicators (RSEI) model, which uses a database from the Environmental Protection Agency of toxic release emissions to air, water, and land, and 2) chemical mixtures measured in house dust (groups of PCBs, PAHs, and two mixtures of pesticides) for study participants enrolled in the NCI-SEER population-based case-control study (1998–2000) at four SEER centers – Detroit, Iowa, Los Angeles County, and Seattle. We assigned 11 years of annual temporally-varying historic environmental exposure scores by intersecting residential locations from participants' residential histories with a fine grid from the RSEI model and by performing inverse distance weighting between facilities releasing specific carcinogenic chemicals and residential locations for spatially-precise exposure assignments. We used Bayesian index low-rank kriging multiple membership models to identify important lag times for RSEI scores, cumulative effects of RSEI scores, and specific carcinogenic chemical releases into the environment. We found a significant positive association between RSEI scores and NHL at the maximum time lag of 11 years (OR = 1.17, 95% CI (1.06, 1.32)) and a significant cumulative RSEI score effect (OR = 1.30, 95% CI (1.02, 1.84)) for long-term residents in Detroit, where benzene and trichloroethylene were the most important chemicals driving this association. Additionally, we identified significant inverse associations for two study centers and time lags that did not persist in cumulative exposure models. Large weights for dichloromethane and pentachlorophenol in models of cumulative exposure also support evidence for their association with NHL risk. These results underscore the importance of considering historic and cumulative

Corresponding Author: David C. Wheeler, dcwheeler@vcu.edu, Telephone: 804-828-9827, Mailing address: 830 E Main St, Richmond, VA 23298.

environmental exposures and using residential histories for diseases with long latency periods such as NHL.

Keywords

exposome; mixture analysis; environmental exposure; non-Hodgkin lymphoma; spatial analysis

Introduction

The exposome is an emerging framework in epidemiological research. Specifically, this approach posits that individuals experience risk of developing adverse health outcomes from a variety of exposure sources across their life course[1,2] and consists of internal, specific external, and general external domains. Many mixtures of exposures act on individuals simultaneously, including stressors from the natural, social, policy, and built environments[3], and the effects of some of these mixtures can increase the likelihood of disease over time. The utility of exposome-driven research is particularly apparent for cancer etiology, as many cancers can have a long latency period between exposure and diagnosis[4–6].

One developing class of models for health outcomes estimates the effects of mixtures of exposures, such as chemical concentrations measured in biomarkers or house dust. In contrast to single-chemical analyses, these models estimate an overall mixture effect of a group or several groups of chemicals, which more realistically capture the effects of many chemical exposures that act on individuals at once[7]. For example, one analysis used Bayesian kernel machine regression models to estimate the effects of correlated groups of exposures to toxic metals collected in urine samples on oxidative stress biomarkers[8]. Another analysis simultaneously estimated the effects of four groups of chemical exposures measured in house dust on risk for non-Hodgkin lymphoma (NHL) using a Bayesian index low-rank kriging multiple membership model[9]. These and other models have generally focused on exposures measured inside the body or home and have often identified associations between them and disease risk[8–11].

In addition to mixtures of chemicals measured in the home or biomarkers, exposure to environmental chemicals near the home are of interest, particularly for exposures farther back in time when disease latency is long. For example, a feasibility study identified residences that were likely to be exposed to common agricultural pesticides by combining satellite imagery, study participants' residential histories, and geographic information system (GIS) methods[12]. An analysis in California also used GIS methods to link pesticide use reporting and crop location data from separate state-administered databases in order to predict the existence of pesticide use at residential locations[13]. Retrospective analyses of NHL have identified elevated risk associated with residential proximity to dioxin-emitting facilities such as lumber and wood products[14] and cement kilns[15]. A commonality of these analyses is defining exposure based on proximity to a putative disease-causing mechanism and assessing its association with disease status. An alternative strategy is to use a comprehensive metric that incorporates many disease mechanisms into one score for exposure assignment. The Environmental Protection Agency (EPA) produces summary

metrics through its Risk-Screening Environmental Indicators (RSEI) model[16], which applies toxicity weights that take into account carcinogenicity and potential human exposure factors to releases from the Toxics Release Inventory (TRI) database[17] in order to produce a unitless and relative measure of environmental toxicity score for small areas across the United States. For example, the TRI provides estimates of emissions to air for chemicals such as benzene, which is a known carcinogen for NHL[18–20]. The estimated emissions for this and other chemicals are combined into an overall score which can be used for an association with disease risk. This score provides an opportunity to assess the degree of potential historical exposure to toxic chemicals in the absence of measurements on the individual level. Additionally, the EPA produces RSEI scores specific to individual chemical releases, as well as carcinogenic and non-carcinogenic sub-scores.

We used historic environmental risk scores for residential histories from the National Cancer Institute (NCI) Surveillance, Epidemiology, and End Results (SEER) case-control study of NHL to address the specific external component of the exposome and NHL risk. Previous spatial analyses of residential history data from this study have found areas of significantly elevated unexplained risk for NHL in three of four study centers (Detroit, Iowa, and Los Angeles) that persisted after adjustment for several environmental, demographic, and genetic risk factors collected after diagnosis[21–23]. A goal of this analysis was to determine whether historic environmental exposures based on residential histories are associated with risk for NHL. To do this, we modeled NHL and 11 years of historic carcinogenic RSEI environmental risk scores using residential histories while adjusting for chemical mixtures measured inside the home. For significant associations between historic environmental risk scores and NHL, we modeled an index of chemical-specific RSEI scores for NHL in order to assess the role of specific chemical releases in NHL risk.

Methods

Study Population.

The NCI-SEER NHL study is a population-based, multi-center case-control study of NHL in four areas of the United States (Wayne, Oakland, and Macomb Counties, comprising the Detroit metropolitan region; the state of Iowa; Los Angeles County; and King and Snohomish Counties, comprising the Seattle metropolitan region). The study population, which has been described in detail previously[24,25], included 1,321 cases of NHL aged 20 to 74 years, who were diagnosed between July 1, 1998 and June 30, 2000, at one of the above four SEER registries. Population-based controls (1,057) were selected among the residents of each SEER registry using random-digit dialing for controls younger than 65 years old and Medicare eligibility files for controls older than 65 years. Eligible controls (who did not have a history of NHL or HIV) were frequency matched to the cases by age within 5-year groups, sex, race, and study center. Table 1 summarizes the characteristics of the study population by study center.

Study participants completed a lifetime residential history calendar, which asked participants to state the complete address of any home they lived in, beginning from birth and including temporary or vacation homes where they lived for a total of at least two years. Interviewers completed in-person interviews with participants in which they reviewed the calendar with

participants and attempted to resolve any discrepancies or missing data in the calendar. Residential addresses were geocoded by matching to street databases in a geographic information system to obtain geographic coordinates[26]. Interviewers took global positioning system (GPS) readings outside the home to obtain the coordinates for the current home. Previous analyses of this data have reported high rates of geocoded addresses and high address matching quality[23], and the mean number of years in residential histories before study entry among participants in our sample was between 52 and 60 years at all study centers.

Vacuum cleaner dust was sampled from the homes of consenting participants, the details of which have been described elsewhere[27,28]. The following chemicals were measured using gas chromatography/mass spectrometry as previously[28] described: polychlorinated biphenyls (PCBs) (congeners 105, 138, 153, 170, 180); polycyclic aromatic hydrocarbons (PAHs) (benz(a)anthracene, benzo(a)pyrene, benzo(b)fluoranthene, benzo(k)fluoranthene, chrysene, dibenz(ah)anthracene, indeno(1,2,3-cd)pyrene); pesticides (group I) (α-chlordane, γ-chlordane, carbaryl, dichlorodiphenyldichloroethylene (DDE), dichlorodiphenyltrichloroethane (DDT), *o*-phenylphenol, pentachlorophenol, propoxur); and pesticides (group II) (chlorpyrifos, *cis*-permethrin, *trans*-permethrin, 2,4-D, diazinon, dicamba, methoxychlor). We chose this grouping of chemicals for the chemical mixture analysis based on previous chemical analysis of NHL risk with univariate associations[29,30]. The pesticides were split into two groups according to positive and negative univariate association with NHL. Some measurements contained missing values, which generally occurred when the chemical concentration fell below the detection limit of the measuring instrument. Assuming a log-normal distribution for concentrations, multiple imputation was used to replace the missing values in the range of 0 and the detection limit in ten datasets with a fill-in approach, the details of which have been described previously[27,29,31]. We used one imputed dataset at random in our analysis.

Historic industrial emissions.

The United States EPA requires that releases of specific chemicals into the environment from federal and industrial facilities be reported through the Toxic Release Inventory (TRI). This program was authorized by an act of Congress in 1986, and annual reporting of chemical releases began in the following year[17]. TRI records the location of the emitting facility and quantity in volumes of over 700 chemicals released above a reporting threshold volume annually to air, water, and land both on- and off-site that may cause chronic health effects or significant acute or environmental health effects. The RSEI model[16] developed by EPA integrates information about the chemical's mode of transport, toxicity, and potential for human exposure with the TRI data to produce a unitless score of relative environmental toxicity. The carcinogenic RSEI score is available over a grid covering the United States, where each cell has size 810 meters by 810 meters. Additionally, the RSEI model produces scores associated with chemical releases at each facility. To identify specific chemicals associated with NHL risk, we used the International Agency for Research on Cancer (IARC) monographs[20] to identify chemicals with sufficient or limited evidence to cause NHL in humans. These chemicals spanned Groups 1 to 2B ("The agent is carcinogenic to humans" to "The agent is possibly carcinogenic to humans") defined by IARC.

To determine annual environmental exposure risk for study participants, we spatially intersected the annual residential locations from participant residential histories with the RSEI grid and assigned the annual carcinogenic RSEI scores to participants for the period 1988–2000. For all models fitted in the form of model 1 and model 2, we fit separate models assigning exposure from 1) the carcinogenic RSEI score in the grid cell where the participant lived or 2) the average of the carcinogenic RSEI scores in that grid cell and its first-order neighbor grid cells based on queen contiguity. We assigned exposure based on adjacency to represent activity spaces and a larger definition of neighborhoods for participants beyond the immediate residential area. This could also address potential edge effects for residences near the border of the grid cell containing it. For models fitted in the form of model 3, we also assigned annual exposures of the chemical-specific RSEI scores using inverse distance weighting between the residential location of the participant at that time and all the locations of the facilities releasing chemicals in the state containing the residential location. In the statistical analysis described below, we used the logarithm of RSEI scores, owing to considerable positive skew in these quantities.

Model Specification.

We used a Bayesian index low-rank kriging multiple membership model (LRK-MMM) to model the probability that a study participant had NHL at each SEER study center, treating NHL status as a binary response variable Y taking values of 1 and 0 for cases and controls, respectively. Assuming that the response variable $Y_i \sim \text{Bernoulli}(p_i)$, where p_i represents the probability of case membership, we specified three models according to different forms or functions of the environmental exposure risk scores:

- Model 1: Carcinogenic RSEI score at time lag t, $t = 0, \ldots, 11$ years before study entry

$$
\begin{aligned}
log\left(\frac{p_i}{1-p_i}\right) = {} & \beta_0 + \beta_1 \sum_{j=1}^{J_1} w_{j1} q_{ij1} + \beta_2 \sum_{j=1}^{J_2} w_{j2} q_{ij2} + \beta_3 \sum_{j=1}^{J_3} w_{j3} q_{ij3} + \beta_4 \sum_{j=1}^{J_4} w_{j4} q_{ij4} + \beta_5 * rsei_{it} \\
& + \sum_{b=1}^{B} \theta_b x_{ib} + \sum_{r \in A_{(i)}} \omega_{ir} \sum_{m=1}^{n_\kappa} \psi_m C[|s_{ir} - \kappa_m|/\rho]
\end{aligned}
$$

- Model 2: Cumulative carcinogenic RSEI score over all t, $t = 0, \ldots, 11$ years before study entry

$$
\begin{aligned}
log\left(\frac{p_i}{1-p_i}\right) = {} & \beta_0 + \beta_1 \sum_{j=1}^{J_1} w_{j1} q_{ij1} + \beta_2 \sum_{j=1}^{J_2} w_{j2} q_{ij2} + \beta_3 \sum_{j=1}^{J_3} w_{j3} q_{ij3} + \beta_4 \sum_{j=1}^{J_4} w_{j4} q_{ij4} \\
& + \beta_5 \sum_{t=0}^{11} w_t rsei_{it} + \sum_{b=1}^{B} \theta_b x_{ib} + \sum_{r \in A_{(i)}} \omega_{ir} \sum_{m=1}^{n_\kappa} \psi_m C[|s_{ir} - \kappa_m|/\rho]
\end{aligned}
$$

- Model 3: Cumulative chemical-specific carcinogenic RSEI index

$$log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 \sum_{j=1}^{J_1} w_{j1} q_{ij1} + \beta_2 \sum_{j=1}^{J_2} w_{j2} q_{ij2} + \beta_3 \sum_{j=1}^{J_3} w_{j3} q_{ij3} + \beta_4 \sum_{j=1}^{J_4} w_{j4} q_{ij4}$$

$$+ \beta_5 \sum_{t=0}^{11} \gamma_t \sum_{j=1}^{J_5} w_{jt} q_{ijt} + \sum_{b=1}^{B} \theta_b x_{ib} + \sum_{r \in A_i} \omega_{ir} \sum_{m=1}^{n_K} \psi_m C[|s_{ir} - \kappa_m|/\rho]$$

There are several components that are common to all models. Specifically, β_0 is an intercept term, and $[\beta_1, \ldots, \beta_4]$ are the estimated health effects associated with the exposure indices for PCBs, PAHs, Pesticides I, and Pesticides II groups of chemicals measured at one time inside the homes of participants. We adjusted for these indices in all models owing to their associations with NHL risk in a previous analysis[9] and to permit interpretation of the historic industrial emissions index effects in the presence of chemicals measured inside the home. For these indices, the estimated importance weight vectors $[w_1, w_2, w_3, w_4]$ are defined such that the weights $w_{jk} \in (0,1)$ for all k in the jth index and $\sum_k w_{jk} = 1$. Additionally, the term q_{ijk} represents the jth exposure in the kth index for the ith individual. Also, the coefficients for the adjustment covariates $[x_{i1}, \ldots, x_{iB}]$ are given by $[\theta_1, \ldots, \theta_B]$. Finally, the spatial risk component of the model is designed to estimate the cumulative spatial risk incurred by participants over their residential histories. The set $[\kappa_1, \ldots, \kappa_{n_\kappa}]$ represents the n_κ knot locations where the spatial random effects are estimated (further detail is given below). The residential history for the ith subject is given by the set of locations $A_{(i)} = [s_{i1}, \ldots, s_{iR}]$, with proportions of the time the subject lived in these locations $[\omega_{i1}, \ldots, \omega_{iR}]$. The set of spatial random effects is given by $[\psi_1, \ldots, \psi_{n_K}]$ where each element of ψ is evaluated at a knot location chosen by the knot selection algorithm.

The other components of the model vary for the different models of environmental exposure. In model 1, the coefficient β_5 represents the health effect associated with the ith participant's exposure to $rsei_{it}$, the carcinogenic RSEI score at time lag t, $t = 0, \ldots, 11$ years before study entry. We chose the maximum time lag $t = 11$ based on the years of diagnosis/reference date for cases and control participants and the first year of reporting for TRI (1988). Model 2 assumed a more comprehensive effect of environmental exposure and allowed the relative importance of different time lags for NHL risk to vary. Here, the coefficient β_5 represents the health effect associated with the cumulative carcinogenic RSEI score exposure index $\sum_{t=0}^{11} w_t rsei_{it}$, with estimated importance weights for the yearly importance weights $[w_0, \ldots, w_{11}]$. For these models, we also fit separate models using data from 1) all participants within a study center or 2) only residents whose entire residential history was contained within the geographic bounds of a study center (i.e., long-term residents of the study center). We fit models in the form of model 3 to explore the role of specific chemicals' RSEI scores over time for NHL risk in long-term residents of a study center. Our rationale in doing so was to identify the specific chemicals that may drive significant associations between an overall carcinogenic RSEI score exposure and risk for NHL. For model 3, β_5 represents NHL risk associated with the cumulative carcinogenic chemical-specific RSEI index, and γ_t represents the importance weight for year lag t. Here, we included all J_5 chemicals in the environmental exposure index that have been classified as having sufficient or limited evidence to cause NHL in humans by IARC and appeared in the TRI at the appropriate geographic and temporal dimensions. Though the presence of these chemicals in the TRI varied by study

center and year (Supplemental Material Table S5), the index at a given center was a subset of the following chemicals (lindane, pentachlorophenol, benzene, dichloromethane, ethylene oxide, trichloroethylene). Additionally, for the chemical importance weight vector w_{jt}, $t = 0, \ldots, 11$, we used the vector of mean estimated weights for chemicals in the same environmental exposure index estimated for separate models of RSEI exposures from lag t, $t = 0, \ldots, 11$.

We fitted Bayesian index LRK-MMMs at each study center to allow for differences in chemical and environmental exposure profiles and strengths of association with NHL status in different regions of the country. We adjusted for age, gender (male vs. reference female), race (Black or Other vs. reference White), and level of education (college degree or high school degree vs. reference less than high school degree) in all models, as in previous analyses of data from this study[21,25,28].

Model Fitting.

To address the spatial risk component, we used $n_K = 60$ knots in all models, which provided an adequate balance of representing the spatial distribution of participants and managing computational costs. We used the Teitz and Bart heuristic[32] to choose knot locations for all models described below due to its demonstrated ability for good spatial sensitivity and power to identify regions of elevated spatial risk in case-control studies[33,34].

For prior distributions, all regression parameters β_i received a Normal prior $\beta_i \sim N(0, \tau_i)$, where the precision term $\tau_i = \frac{1}{\sigma_i^2}$ and $\sigma_i \sim Unif(0,10)$. The spatial random effects ψ received a multivariate normal prior $MNV(0, \tau_s\Omega^{-1})$, where $\tau_s = \frac{1}{\sigma_s^2}$ and $\sigma_s \sim Unif(0,10)$, and the precision matrix $\Omega = [C[|k_a - k_b|/\rho]]$, $1 \le a, b \le n_K$. We used the Matern family for the spatial covariance function, which simplifies to $C[|f|] = (1 + |f|/\rho) \, exp\,(-|f|/\rho)$ for distance f when fixing parameters in the family of m to 1 and ν to $\frac{3}{2}$, respectively. This family of covariance functions has been used in geostatistical models due to its flexibility and smoothness[35,36]. The spatial correlation parameter ρ was assigned a vague $Unif(0, \rho_{max})$ prior, where ρ_{max} is the maximum distance between a knot location and a cell in a fine grid covering the study region. The importance weights in each index received a Dirichlet prior with parameter vector $\alpha = [\alpha_1, \ldots, \alpha_{C_j}]$, where C_j denotes the number of components in the index, in order that the importance weights were between 0 and 1 and $\sum_{j=1}^{C_j} W_{ij} = 1$.

We estimated models with Just Another Gibbs Sampler (JAGS)[37] in the software R[38], version 3.6.1, using Markov chain Monte Carlo (MCMC) methods. In the MCMC simulations, each model used two chains that each burned in 80,000 iterations and sampled 80,000 observations from the posterior distribution. We assessed convergence of model parameters using the Gelman-Rubin statistic, considering a parameter to have converged if its statistic was less than 1.1[39], using the coda[40] R package.

Results

Scores for exposure assignments varied considerably by study center and generally decreased over the study period (Supplemental Material tables S1–S4). RSEI scores were higher in Detroit and Los Angeles than in Iowa and Seattle, with the highest scores at a time lag of 11 years in Los Angeles and Detroit. Additionally, assigning exposure based on first-order neighbors often led to lower RSEI scores than did assignment based on the grid cell where the participant lived. The posterior means and 95% credible intervals for the odds ratios from model 1 (Figure 1) show some significant associations between historic environmental exposures as measured by carcinogenic RSEI scores at specific time lags and NHL risk after adjusting for chemicals measured in the home, individual covariates, and cumulative spatial risk.

When using the exposure from the residential grid cell (panel (a)), there exists considerable variability in the effect estimates at the Detroit, Los Angeles, and Seattle study centers. Notably, there is a significant and positive association between RSEI score exposure and risk for NHL in Detroit at 11 years before study entry (OR = 1.07, 95% CI (1.02, 1.13)) that increases in strength when only considering long-term residents (OR = 1.17, 95% CI (1.06, 1.32)). Additionally, there is a slight inverse association for RSEI score for all participants in Los Angeles at lag 11 years (OR = 0.91, 95% CI (0.86, 0.96)), and among long-term residents in Seattle at lag 9 years (OR = 0.67, 95% CI (0.43, 0.97)) and Los Angeles at lag 11 years (OR = 0.84, 95% CI (0.69, 0.96)). Generally, there are positive if not significant odds ratios at most time lags in Detroit, Iowa, and among all residents in Seattle. Averaging the RSEI scores over first-order neighbor grid cells increased precision in the RSEI effect estimates (panel (b)), as the credible intervals became notably narrower in Detroit, Los Angeles, and Seattle. Additionally, the RSEI score odds ratios became more positive for many year lags in Los Angeles and the significant inverse association at lag 11 years did not remain. Performing exposure assignment in this way, the significant positive associations at lag 11 years in Detroit remained for all participants (OR = 1.05, 95% CI (1.00, 1.10)) and long-term residents (OR = 1.09, 95% CI (1.00, 1.19)). There were additional inverse associations at lag 11 years in Iowa (OR = 0.94, 95% CI (0.89, 0.99)) and among long-term residents in Seattle (OR = 0.89, 95% CI (0.79, 0.98)).

In the model 2 results, there was a significant and positive association between cumulative RSEI score exposure and NHL risk among long-term residents in Detroit when using the residential grid cell (OR = 1.29, 95% CI (1.00, 1.82)) and all first-order neighbor (OR = 1.30, 95% CI (1.02, 1.84)) grid cells for exposure assignment (Table 2). In these models, exposure at 11 years before study entry received the largest estimated importance weight (0.48 and 0.47, respectively), which agreed with the significant and positive association for the single-year exposure at this time and study center in Model 1. Additionally, 12 of the 16 cumulative RSEI score models in the table had estimated importance weights for lag 11 of greater than 0.10, and in 10 of the 16 models the lag 11 weight received the largest estimated weight over all years. In Iowa, the cumulative RSEI score exposure index had an elevated but not significant odds ratio for all combinations of grid cells used for exposure assignment and inclusion of only long-term residents. In Seattle, the cumulative RSEI index had an elevated but not significant odds ratio for all residents using both methods of

grid cell-based exposure assignment. We emphasize that while both of these indices have elevated associations with NHL risk, they are not significant in these models.

In all study centers, the cumulative chemical-specific carcinogenic RSEI index in Model 3 was positively associated with NHL risk (Table 3), particularly in Detroit (OR = 1.35, 95% CI (0.89, 2.61)). The yearly weights in this cumulative index illustrate the importance of historic environmental exposures, with elevated weights at lag 10 and 11 years in Detroit, lags 7, 8, and 11 in Iowa (index OR = 1.17, 95% CI (0.56, 2.86), lag 11 in Los Angeles (index OR = 1.11, 95% CI (0.61, 2.12), and lag 10 in Seattle (index OR = 1.16, 95% CI (0.74, 2.03) (Figure 2). We note that despite elevated associations with NHL risk at each study center, the cumulative chemical-specific carcinogenic RSEI index was not significantly associated with NHL risk at any study center.

Over this time range, the chemical importance weights tended to be largest for pentachlorophenol in Iowa, Los Angeles, and Seattle, and for dichloromethane in Detroit (Figure 3). In addition, large chemical importance weights in indices close to study entry may be related to weights for these chemicals at longer lag times due to possible similarities in chemical emissions profiles over space and time.

At each center, the annual models that estimated these chemical importance weights (Supplemental Material Table S6) had improved goodness of fit at certain time lags compared to the cumulative model. For example, in Detroit at lag 11, the chemical-specific RSEI index was significantly and positively associated with NHL (OR = 1.11, 95% CI (1.00, 1.25)), and this model had lower DIC than the cumulative one. In this index, benzene (0.327) and trichloroethylene (0.266) received a majority of the importance weight.

Discussion

In this study, we estimated the associations between historic environmental pollutant exposure measures and NHL risk using Bayesian index LRK-MMMs, which use residential histories to adjust for cumulative spatial risk when estimating mixture effects. We found a significant positive association between carcinogenic RSEI scores at a time lag of 11 years before diagnosis and NHL risk for residents of the Detroit study area. Eleven years was the farthest back in time we could go based on availability of RSEI data. Notably, these associations did not persist at times closer to diagnosis. We also identified significant inverse associations with NHL risk among participants in Los Angeles at lag 11 years and in Seattle at lag 9 years. This finding may be attributable to a number of factors. For example, because the estimated association in Los Angeles at lag 11 years is markedly different than its counterparts in all other years, it is possible that higher variability in chemical emissions estimates to the TRI in its first year of existence may have led to estimated associations that were not borne out in later years. Further, we found a significant positive association between cumulative carcinogenic RSEI scores over the 11-year exposure period and NHL for long-term residents in Detroit, with exposure scores at 11 years having the largest importance weight in the index. Using a cumulative exposure index also attenuated the significant inverse associations in Los Angeles and Seattle at lags 11 and 9 years respectively to the null. Following up on these findings, we identified

positive associations between NHL and an index of chemical RSEI scores among long-term residents at each study center, with pentachlorophenol and dichloromethane showing the strongest associations with risk. Additionally, we identified a significant and positive association for an index of specific carcinogens from RSEI in Detroit at lag 11, with benzene and trichloroethylene receiving the majority of the weight in the index. This significant association could explain some areas of elevated risk for NHL in Detroit from previous analyses[22,23] that persisted after adjusting for environmental, demographic, occupational, and genetic factors.

The findings in this study contribute to the existing literature surrounding environmental risk factors and NHL. Some studies have identified elevated risk for NHL from exposures to groups of organic solvents[18,41–43] and proximity to facilities in the lumber, wood products, and chemical industries[14]. However, exposures to specific chemicals such as benzene[18–20,44,45], diazinon[20,46], DDE[47], polychlorinated dibenzo-p-dioxins and dibenzofurans[48], toluene[18,19], styrene[19], PCB congeners[20,47,49–51], and trichloroethylene[20,44,52,53] have demonstrated varying degrees of association with NHL risk. Our study sought to add to the literature in several ways. First, our modeling strategy used an extension of the Bayesian group index model, which has demonstrated the ability to accurately estimate mixture effects and the importance of components in mixtures in the presence of high within- and between-mixture correlations[54], as well as accommodate different directions of association between several mixtures and the outcome simultaneously. Second, we used a large public database with fine spatial precision to estimate environmental exposure measures based on all reported chemical releases as well as environmental toxicity weights. Third, we used residential histories to consider historic environmental exposures. Through these analytic choices, our study was able to estimate a series of annual environmental exposure risk effects in a case-control study that adds evidence to the literature for the associations of benzene and trichloroethylene with NHL risk due to the significant annual finding in Detroit. Additionally, the large weights across many years in the cumulative models for pentachlorophenol and dichloromethane support evidence for their association with NHL risk. Notably, the evidence for pentachlorophenol carcinogenicity in humans is considered by IARC to be sufficient; whereas, evidence for benzene, trichloroethylene, and dichloromethane and NHL is considered limited[20]. The EPA recently prohibited the manufacture, sale, and distribution of pentachlorophenol in the United States[55].

The strengths of our study should be considered alongside its limitations. First, the EPA's calculation of RSEI scores relies upon the chemical releases reported in the TRI database, which may be incomplete or reported with error in released chemical volumes due to industrial reporting practices. Second, while the many null associations estimated in our study are not inherently a limitation, they may demonstrate the challenges in working with historic chemical emissions data, particularly at individual time lags. This may encourage the use of cumulative and more stable environmental exposure indices in future research. Additionally, the dates of enrollment for the NCI-SEER NHL study and of the beginning of reporting to the TRI enabled us to estimate environmental exposures at a maximum of only 11 years before study entry for all participants. The high estimated importance weights of exposures at the longest lag – correspondingly, generally when exposure scores

were higher – suggest that stronger associations between environmental exposures and NHL risk exist farther back in time. We only included chemicals in the index for long-term residents in Model 3 that were IARC-identified carcinogens for NHL that appeared in this time range in the database of chemical releases (Supplemental Material Table S5). The lack of reported chemical emissions to TRI at other study areas at this time point (e.g., there were no reported releases of pentachlorophenol in Detroit) may indicate that high exposures to specific carcinogens for NHL drove the significant associations found in Detroit, and not simply high carcinogenic RSEI scores. Also, our analysis did not distinguish between different histologic subtypes of NHL and treated all subtypes identically. Because considerable etiologic heterogeneity has been demonstrated for different NHL subtypes, particularly in risk factors between T-cell and B-cell lymphomas[56], this work was unable to capture how different environmental exposures may be associated with risk of different NHL subtypes. Finally, non-participation was present in the formation of our analysis sample. Some eligible cases and controls were not included in the sample because they had died, they or their physician refused, they could not be located, they moved out of the study area, or they were otherwise ill, impaired, or unable to be contacted. However, a spatial analysis of NHL risk did not find a substantial change in risk estimates between all eligible participants and actual participants[23]. Additionally, an analysis of the bias regarding educational attainment in census tracts between participants and eligible non-participants found only a small degree of bias (between 1% and 8%) in the estimates for NHL risk[57].

In summary, our study findings demonstrate the importance of considering historic environmental exposure risk in the context of diseases with long latency periods such as NHL. Future studies focusing on the effects of environmental exposures should consider using a long period of historic exposures, which is becomingly increasingly possible because reporting to the TRI continues into the present day. In addition, future studies should leverage existing and novel methods to estimate the effects of environmental exposures to chemicals such as benzene and trichloroethylene on NHL risk. Finally, similar studies to ours in other geographic areas will help to guide policies that limit the extent of chemical releases into the environment, with a focus on releases into low-socioeconomic status neighborhoods, which often bear the brunt of environmental chemical exposures[58].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

References

1. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Prev Biomarkers. 2005;14(8):1847–1850.

2. Wild CP. The exposome: from concept to utility. Int J Epidemiol. 2012;41(1):24–32. [PubMed: 22296988]

3. Juarez PD, Hood DB, Song M-A, Ramesh A. Use of an exposome approach to understand the effects of exposures from the natural, built, and social environments on cardio-vascular disease onset, progression, and outcomes. Front public Heal. 2020;8:379.

4. Archer VE, Coons T, Saccomanno G, Hong D-Y. Latency and the Lung Cancer Epidemic Among United States Uranium Miners. Health Phys. 2004;87(5). https://journals.lww.com/health-physics/Fulltext/2004/11000/LATENCY_AND_THE_LUNG_CANCER_EPIDEMIC_AMONG_UNITED.4.aspx

5. Miyakawa M, Tachibana M, Miyakawa A, et al. Re-evaluation of the latent period of bladder cancer in dyestuff-plant workers in Japan. Int J Urol. 2001;8(8):423–430. [PubMed: 11555006]

6. Smith AH, Lopipero P. Invited commentary: How do the Seveso findings affect conclusions concerning TCDD as a human carcinogen? Am J Epidemiol. 2001;153(11):1045–1047. [PubMed: 11390320]

7. de Vuijst E, van Ham M, Kleinhans R. A Life Course Approach to Understanding Neighbourhood Effects. IZA Discussion paper #10276; 2016.

8. Kim SS, Meeker JD, Keil AP, et al. Exposure to 17 trace metals in pregnancy and associations with urinary oxidative stress biomarkers. Environ Res. 2019;179:108854. [PubMed: 31678726]

9. Boyle J, Ward MH, Cerhan JR, Rothman N, Wheeler DC. Estimating mixture effects and cumulative spatial risk over time simultaneously using a Bayesian index low-rank kriging multiple membership model. Stat Med. Published online 2022.

10. Wheeler DC, Rustom S, Carli M, Whitehead TP, Ward MH, Metayer C. Assessment of grouped weighted quantile sum regression for modeling chemical mixtures and cancer risk. Int J Environ Res Public Health. 2021;18(2):504. [PubMed: 33435473]

11. Midya V, Colicino E, Conti DV, et al. Association of Prenatal Exposure to Endocrine-Disrupting Chemicals With Liver Injury in Children. JAMA Netw open. 2022;5(7):e2220176--e2220176. [PubMed: 35793087]

12. Ward MH, Nuckols JR, Weigel SJ, Maxwell SK, Cantor KP, Miller RS. Identifying populations potentially exposed to agricultural pesticides using remote sensing and a Geographic Information System. Environ Health Perspect. 2000;108(1):5–12.

13. Nuckols JR, Gunier RB, Riggs P, Miller R, Reynolds P, Ward MH. Linkage of the California Pesticide Use Reporting Database with spatial land use data for exposure assessment. Environ Health Perspect. 2007;115(5):684–689. [PubMed: 17520053]

14. De Roos AJ, Davis S, Colt JS, et al. Residential proximity to industrial facilities and risk of non-Hodgkin lymphoma. Environ Res. 2010;110(1):70–78. [PubMed: 19840879]

15. Pronk A, Nuckols JR, De Roos AJ, et al. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case-control study. Environ Heal. 2013;12(1):1–11.

16. EPA. Risk-Screening Environmental Indicators (RSEI) Model. Published 2022. https://www.epa.gov/rsei

17. EPA. Toxics Release Inventory (TRI) Program. Published 2022. https://www.epa.gov/toxics-release-inventory-tri-program

18. Vineis P, Miligi L, Costantini AS, et al. Occupational exposure to solvents and the risk of lymphomas. Cancer Epidemiol biomarkers Prev. 2006;16(3):381–384.

19. 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(5):341–347. [PubMed: 20447988]

20. IARC. Agents classified by the IARC monographs, volumes 1–132. International agency for research on cancer. Published 2022. Accessed October 26, 2022. http://monographs.iarc.fr/ENG/Classification/index.php

21. Wheeler DC, Waller LA, Cozen W, Ward MH. Spatial-temporal analysis of non-Hodgkin lymphoma risk using multiple residential locations. Spat Spatiotemporal Epidemiol. 2012;3(2):163–171. [PubMed: 22682442]

22. Wheeler DC, Calder CA. Sociospatial epidemiology: residential history analysis. In: Handbook of Spatial Epidemiology. Chapman and Hall/CRC; 2016:627–648.

23. 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 Heal. 2011;10(1):1–13.

24. Chatterjee N, Hartge P, Cerhan JR, et al. Risk of non-Hodgkin's lymphoma and family history of lymphatic, hematologic, and other cancers. Cancer Epidemiol Prev Biomarkers. 2004;13(9):1415–1421.

25. Morton LM, Wang SS, Cozen W, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes. Blood, J Am Soc Hematol. 2008;112(13):5150–5160.

26. ESRI. ArcView 3.2

27. Colt JS, Lubin J, Camann D, et al. Comparison of pesticide levels in carpet dust and selfreported pest treatment practices in four US sites. J Expo Sci Environ Epidemiol. 2004;14(1):74–83.

28. Colt JS, Severson RK, Lubin J, et al. Organochlorines in carpet dust and non-Hodgkin lymphoma. Epidemiology. Published online 2005:516–525. [PubMed: 15951670]

29. Czarnota J, Gennings C, Colt JS, et al. Analysis of environmental chemical mixtures and non-Hodgkin lymphoma risk in the NCI-SEER NHL study. Environ Health Perspect. 2015;123(10):965–970. [PubMed: 25748701]

30. Wheeler D, Czarnota J. Modeling Chemical Mixture Effects with Grouped Weighted Quantile Sum Regression. ISEE Conference Abstracts; 2016.

31. Lubin JH, Colt JS, Camann D, et al. Epidemiologic evaluation of measurement data in the presence of detection limits. Environ Health Perspect. 2004;112(17):1691–1696. [PubMed: 15579415]

32. Teitz MB, Bart P. Heuristic methods for estimating the generalized vertex median of a weighted graph. Oper Res. 1968;16(5):955–961.

33. Boyle J, Wheeler DC. Knot selection for low-rank kriging models of spatial risk in case-control studies. Spat Spatiotemporal Epidemiol. 2022;41:100483. [PubMed: 35691650]

34. Owen SH, Daskin MS. Strategic facility location: A review. Eur J Oper Res. 1998;111(3):423–447.

35. Shaddick G, Zidek JV. A case study in preferential sampling: Long term monitoring of air pollution in the UK. Spat Stat. 2014;9:51–65.

36. Diggle PJ, Tawn JA, Moyeed RA. Model-based geostatistics. J R Stat Soc Ser C (Applied Stat. 1998;47(3):299–350.

37. Plummer M, others. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing. Vol 124.; 2003:1–10.

38. R Core Team, Others. R: A language and environment for statistical computing. Published online 2021.

39. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci. 1992;7(4):457–472.

40. Plummer M, Best N, Cowles K, Vines K. CODA: Convergence Diagnosis and Output Analysis for MCMC. R News. 2006;6(1):7–11. https://www.r-project.org/doc/Rnews/Rnews_2006-1.pdf

41. Dryver E, Brandt L, Kauppinen T, Olsson H. Occupational exposures and non-Hodgkin's lymphoma in Southern Sweden. Int J Occup Environ Health. 2004;10(1):13–21. [PubMed: 15070021]

42. Kato I, Koenig KL, Watanabe-Meserve H, et al. Personal and occupational exposure to organic solvents and risk of non-Hodgkin's lymphoma (NHL) in women (United States). Cancer Causes Control. 2005;16(10):1215–1224. [PubMed: 16215872]

43. Fritschi L, Benke G, Hughes AM, et al. Risk of non-Hodgkin lymphoma associated with occupational exposure to solvents, metals, organic dusts and PCBs (Australia). Cancer Causes Control. 2005;16(5):599–607. [PubMed: 15986116]

44. Alexander DD, Mink PJ, Adami H-O, et al. The non-Hodgkin lymphomas: a review of the epidemiologic literature. Int J cancer. 2007;120(S12):1–39.

45. Rana I, Dahlberg S, Steinmaus C, Zhang L. Benzene exposure and non-Hodgkin lymphoma: a systematic review and meta-analysis of human studies. Lancet Planet Heal. 2021;5(9):e633--e643.

46. De Roos AJ, Schinasi LH, Miligi L, et al. Occupational insecticide exposure and risk of n on-Hodgkin lymphoma: A pooled c ase-control study from the InterLymph Consortium. Int J Cancer. 2021;149(10):1768–1786. [PubMed: 34270795]

47. Engel LS, Laden F, Andersen A, et al. Polychlorinated biphenyl levels in peripheral blood and non-Hodgkin's lymphoma: a report from three cohorts. Cancer Res. 2007;67(11):5545–5552. [PubMed: 17545638]

48. Jones RR, Ward MH, Deziel NC, et al. Residential Proximity to Dioxin-emitting Facilities and Risk of Non-Hodgkin Lymphoma in the NIH-AARP Diet and Health Study. Cancer Epidemiol Biomarkers \& Prev. 2021;30(4):808–809.

49. De Roos AJ, Hartge P, Lubin JH, et al. Persistent organochlorine chemicals in plasma and risk of non-Hodgkin's lymphoma. Cancer Res. 2005;65(23):11214–11226. [PubMed: 16322272]

50. Viel J-F, Floret N, Deconinck E, Focant J-F, De Pauw E, Cahn J-Y. Increased risk of non-Hodgkin lymphoma and serum organochlorine concentrations among neighbors of a municipal solid waste incinerator. Environ Int. 2011;37(2):449–453. [PubMed: 21167603]

51. Bertrand KA, Spiegelman D, Aster JC, et al. Plasma organochlorine levels and risk of non-Hodgkin lymphoma in a cohort of men. Epidemiology. 2010;21(2):172. [PubMed: 20087190]

52. Purdue MP, Bakke B, Stewart P, et al. A case--control study of occupational exposure to trichloroethylene and non-Hodgkin lymphoma. Environ Health Perspect. 2011;119(2):232–238. [PubMed: 21370516]

53. Callahan CL, Stewart PA, Friesen MC, et al. Case-control investigation of occupational exposure to chlorinated solvents and non-Hodgkin's lymphoma. Occup Environ Med. 2018;75(6):415–420. [PubMed: 29588333]

54. Wheeler DC, Rustom S, Carli M, Whitehead TP, Ward MH, Metayer C. Bayesian Group Index Regression for Modeling Chemical Mixtures and Cancer Risk. Int J Environ Res Public Health. 2021;18(7):3486. [PubMed: 33801661]

55. EPA. EPA requires cancellation of pentachlorophenol to protect human health. Published 2022. https://www.epa.gov/pesticides/epa-requires-cancellation-pentachlorophenol-protect-human-health

56. Morton LM, Slager SL, Cerhan JR, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes: the InterLymph non-Hodgkin lymphoma subtypes project. J Natl Cancer Inst Monogr. 2014;2014(48):130–144. [PubMed: 25174034]

57. Shen M, Cozen W, Huang L, et al. Census and geographic differences between respondents and nonrespondents in a case-control study of non-Hodgkin lymphoma. Am J Epidemiol. 2008;167(3):350–361. [PubMed: 17989060]

58. Evans GW, Kantrowitz E. Socioeconomic status and health: the potential role of environmental risk exposure. Annu Rev Public Health. 2002;23(1):303–331. [PubMed: 11910065]

Highlights:

- We assigned 11 years of historic environmental exposure scores based on residence.

- Used Bayesian index low-rank kriging multiple membership models to identify lags.

- Found a significant positive association between exposure scores and NHL.

- Benzene and trichloroethylene were the most important chemicals.

- Results show importance of considering historic environmental exposures.
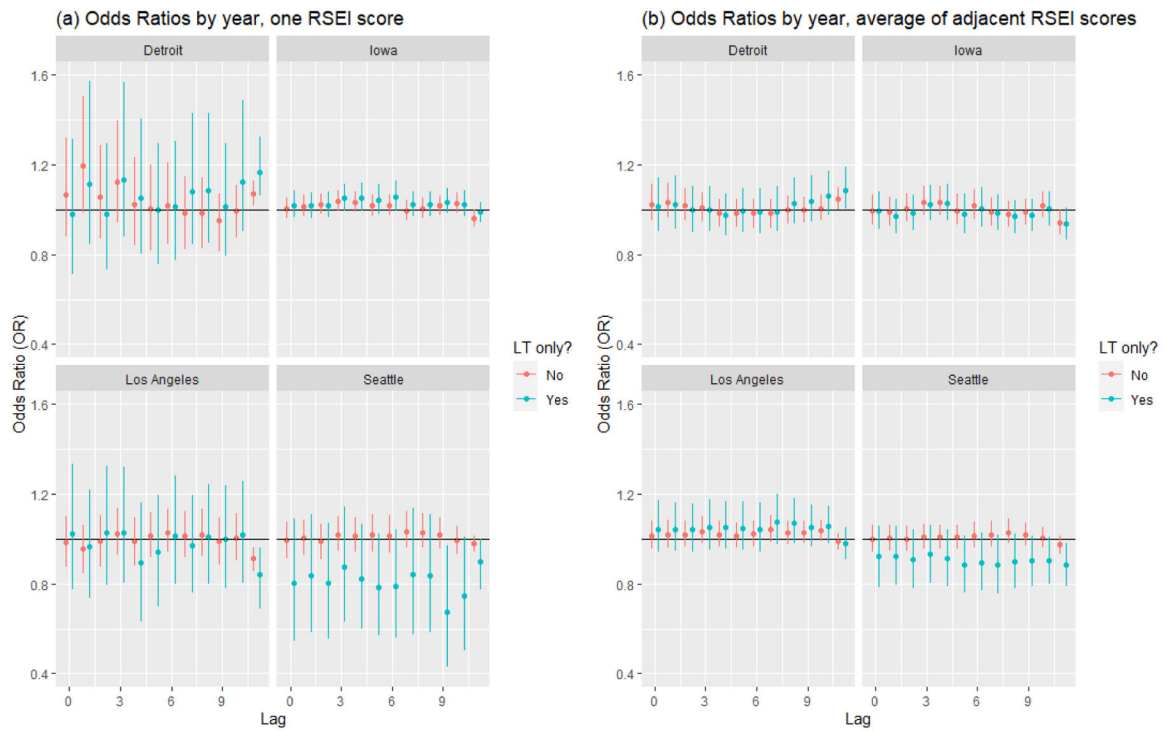
Figure 1.
Summary of carcinogenic RSEI score associations with risk of NHL by time lag and study center from model 1.

Note: Quantities presented in table are posterior mean and posterior 95% credible interval for the odds ratio. Panel (a) assigns RSEI score exposure at one grid cell, and panel (b) assigns RSEI score exposure averaging over all first-order neighbor grid cells, including the one in which the participant lived. All models adjusted for age, gender, race, educational attainment, chemicals measured in the home, and cumulative spatial risk. "LT" denotes long-term residents at a given study center.
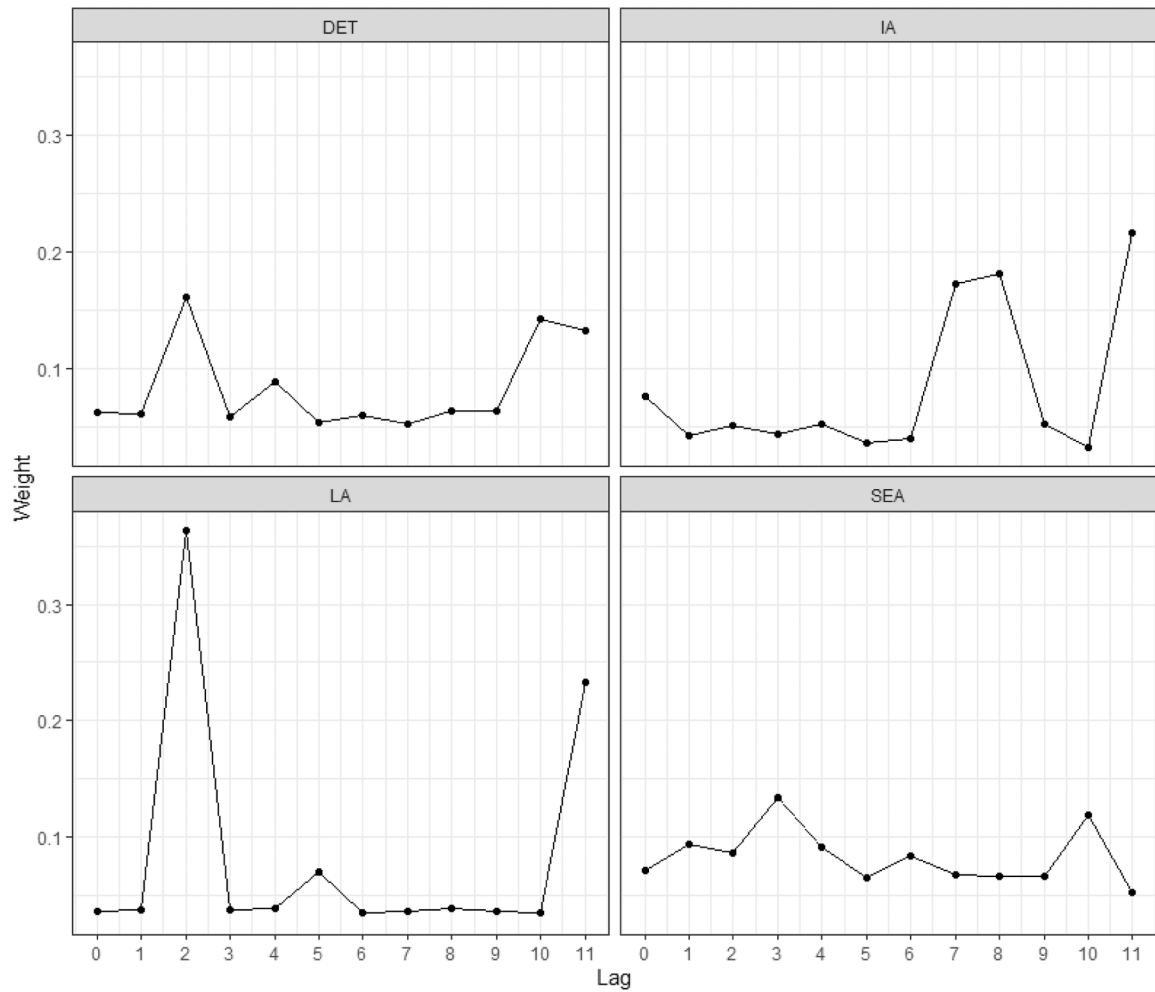
Figure 2.
Summary of estimated yearly importance weights in the cumulative chemical-specific carcinogenic RSEI index in Model 3 by study center. DET = Detroit (Wayne, Oakland, and Macomb counties); IA = Iowa; LA = Los Angeles County; SEA = Seattle (King and Snohomish counties)
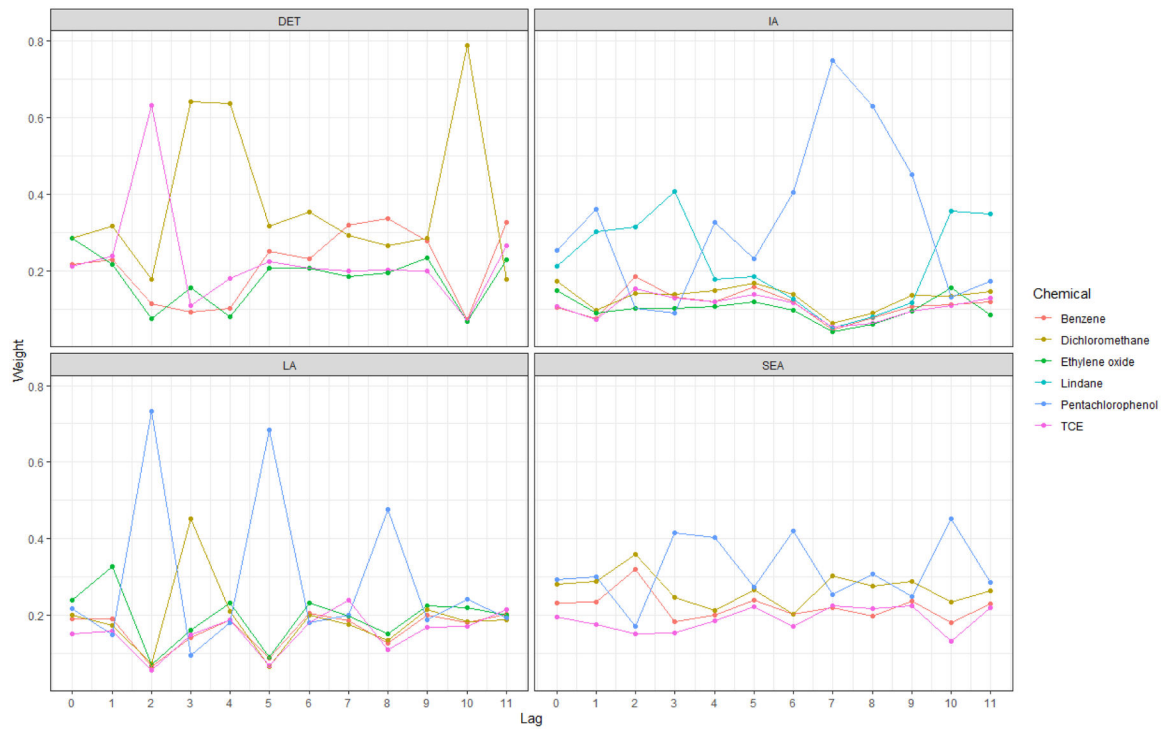
Figure 3.
Summary of estimated chemical importance weights in Model 3 by yearly lag and study center. DET = Detroit (Wayne, Oakland, and Macomb counties); IA = Iowa; LA = Los Angeles County; SEA = Seattle (King and Snohomish counties)

Table 1.

Characteristics of the NCI-SEER NHL study population by study center.

Center	Detroit		Iowa		Los Angeles		Seattle	
Participants	**All**	**LT**	**All**	**LT**	**All**	**LT**	**All**	**LT**
Sample Size	201	99	335	198	292	109	342	81
Status								
Case	127 (63)	64 (65)	188 (56)	106 (54)	168 (58)	63 (58)	182 (53)	47 (58)
Control	74 (37)	35 (35)	147 (44)	92 (46)	124 (42)	46 (42)	160 (47)	34 (42)
Age (years)	58 (11.4)	56 (12.1)	61 (11.2)	61 (11.7)	59 (11.2)	57 (11.4)	59 (10.8)	56 (12.1)
Sex								
Male	114 (57)	54 (55)	177 (53)	105 (53)	165 (57)	63 (58)	171 (50)	45 (56)
Female	87 (43)	45 (45)	158 (47)	93 (47)	127 (43)	46 (42)	171 (50)	36 (44)
Race								
White	164 (81)	83 (84)	331 (99)	195 (98)	215 (74)	73 (67)	316 (92)	75 (93)
Non-white	37 (19)	16 (16)	4 (1)	3 (2)	77 (26)	36 (33)	26 (8)	6 (7)
Education								
< 12 years	23 (11)	12 (12)	32 (10)	23 (12)	31 (11)	15 (14)	19 (6)	8 (10)
12–15 years	124 (62)	67 (68)	241 (72)	152 (77)	171 (59)	67 (61)	201 (59)	48 (59)
>= 16 years	54 (27)	20 (20)	62 (19)	23 (12)	90 (31)	27 (25)	122 (35)	25 (31)

Note: Age summarized using mean (standard deviation) and all other variables summarized using count (percent). "LT" denotes long-term residents at a study center whose entire residential history was contained within the geographic bounds of the study center. Some percentages may not sum exactly to one due to rounding.

Table 2.

Cumulative RSEI score index associations with risk of NHL by study center and model type from model 2.

Center	Grid Cells	Residents	OR	95% CI	0	1	2	3	4	5	6	7	8	9	10	11
			RSEI Score		**Year Lag Importance Weights**											
	One	All	1.12	(0.92, 1.42)	0.07	0.13	0.07	0.08	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.27
	One	LT	1.29	(1.00, 1.82) *	0.04	0.07	0.04	0.05	0.04	0.04	0.04	0.05	0.05	0.04	0.05	0.48
	Avg	All	1.18	(0.98, 1.54)	0.07	0.16	0.07	0.08	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.29
Detroit	Avg	LT	1.30	(1.02, 1.84) *	0.04	0.07	0.04	0.05	0.04	0.04	0.04	0.05	0.06	0.05	0.05	0.47
	One	All	1.01	(0.95, 1.07)	0.07	0.07	0.08	0.11	0.09	0.07	0.08	0.07	0.07	0.07	0.09	0.13
	One	LT	1.04	(0.97, 1.11)	0.07	0.07	0.07	0.11	0.12	0.09	0.11	0.07	0.07	0.09	0.08	0.06
	Avg	All	1.01	(0.95, 1.08)	0.07	0.07	0.08	0.11	0.09	0.07	0.08	0.06	0.07	0.08	0.10	0.13
Iowa	Avg	LT	1.03	(0.97, 1.11)	0.08	0.07	0.07	0.10	0.10	0.08	0.11	0.07	0.07	0.09	0.08	0.06
	One	All	0.91	(0.82, 1.03)	0.03	0.04	0.03	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.69
	One	LT	0.88	(0.67, 1.17)	0.05	0.06	0.05	0.05	0.07	0.06	0.05	0.05	0.05	0.05	0.05	0.40
	Avg	All	0.91	(0.82, 1.03)	0.04	0.04	0.04	0.03	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.63
Los Angeles	Avg	LT	0.86	(0.64, 1.13)	0.06	0.06	0.06	0.06	0.08	0.07	0.05	0.06	0.05	0.05	0.05	0.36
	One	All	1.01	(0.93, 1.10)	0.08	0.08	0.08	0.08	0.08	0.09	0.08	0.09	0.09	0.08	0.08	0.08
	One	LT	0.75	(0.49, 1.03)	0.07	0.07	0.07	0.06	0.07	0.09	0.08	0.07	0.07	0.15	0.10	0.10
	Avg	All	1.01	(0.93, 1.10)	0.08	0.08	0.08	0.08	0.08	0.09	0.09	0.09	0.09	0.08	0.08	0.08
Seattle	Avg	LT	0.79	(0.53, 1.07)	0.07	0.07	0.07	0.06	0.07	0.08	0.08	0.07	0.07	0.14	0.10	0.11

Note: Quantities presented in table are posterior mean and posterior 95% credible interval for the odds ratio. In the "Grid Cells" column, "One" denotes exposure assignment from the residential grid cell, and "Avg" denotes the average of the residential grid cell and its first-order neighbors. In the "Residents" column, "All" denotes all participants from a study area and "LT" denotes only the participants who were long-term residents of the study area. Significant associations, defined as those with credible interval excluding the null odds ratio value of one, are displayed with an asterisk. All models adjusted for age, gender, race, and educational attainment. Model 2 estimated a cumulative carcinogenic RSEI score index.

Table 3.

Odds ratios for cumulative chemical-specific carcinogenic RSEI index among long-term residents.

Center	Odds Ratio	95% CI	DIC
Detroit	1.35	(0.89, 2.61)	238.2
Iowa	1.17	(0.56, 2.86)	422.1
Los Angeles	1.11	(0.61, 2.12)	373.7
Seattle	1.16	(0.74, 2.03)	421.2

Note: DIC represents the deviance information criterion, where lower values indicate better-fitting models.