Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____	_4/21/22___
Rachel Boehm	Date

Environmental and social risk factors in association with spatial clustering of embryonal cancer incidence

By

Rachel Boehm

Master of Public Health

Epidemiology

_____ [Chair's signature]

Anke Hüls, PhD
Committee Chair

Environmental and social risk factors in association with spatial clustering of embryonal cancer incidence

By

Rachel Boehm

Bachelor of Arts
University of Oregon
2013

Thesis Committee Chair: Anke Hüls, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2022

Abstract

Environmental and social risk factors in association with spatial clustering of embryonal cancer incidence

By Rachel Boehm

Co-Authors: Sara Van Cor, Nick Toumbacaris, Liuhua Shi, John L. Pearce, Rana Bayakly, Timothy L. Lash, Kevin Ward, Jeffrey M. Switchenko, Anke Hüls

Cancer is among the top 10 leading causes of death for children aged 0-4 years of age. Embryonal cancers compose roughly 20% of all childhood cancer cases. They almost exclusively appear in young children, with the highest incidence rates, except for ganglioneuroblastoma, in the first year of life. Aside from high-dose radiation and prior chemotherapy, there are no established external risk factors for childhood cancer. The aim of this study is to examine the association of harmful combinations of air pollutants, pesticides and other social risk factors and embryonal cancer incidence among 0-4-year-olds.

This ecological study uses data of 875 embryonal cancer cases registered in the state of Georgia and included those who had a new diagnosis of one of seven subgroups of embryonal cancer. We calculated standardized incidence ratios (SIR) of embryonal cancers based on age, gender and ethnic composition in each of the 159 counties in Georgia, USA. High-quality estimates for $PM_{2.5}$ and its components with a 1 km resolution were used, which are publicly available. Country-level data on pesticides were collected from the USDA National Agricultural Statistics Service searchable database called Quick Stats. Socioeconomic characteristics were measured using 17 county-level indicators corresponding to six socioeconomic domains from the U.S. Census Bureau and the American Community Survey. We applied an unsupervised learning tool (self-organizing map, SOM) to identify pertinent types of multi-exposure combinations. Zero-inflated spatial Bayesian Poisson models (Leroux-CAR) were fit with indicators for each multi-exposure category as exposure and SIR of embryonal cancers as outcomes.

Significantly higher SIRs were observed in areas of increased alcohol consumption. To our knowledge, this is one of the few studies to find a significant association between alcohol consumption and embryonal cancer. Future studies should be conducted to replicate these findings in wider geographic areas and using individual-level data.

Environmental and social risk factors in association with spatial clustering of embryonal cancer incidence

By

Rachel Boehm

Bachelor of Arts
University of Oregon
2013

Thesis Committee Chair: Anke Hüls, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2022

Table of Contents

Introduction ……1

Methods ……2

Results ……4

Discussion……6

References……8

Tables and Figures……12

Introduction

Cancer is among the top 10 leading causes of death for children aged 0-4 years of age [1] . Embryonal cancers, which are tumors composed of undifferentiated cells like that of a developing embryo, compose roughly 20% of all childhood cancer cases[2]. Embryonal cancers almost exclusively appear in young children[2], with the highest incidence rates, except for ganglioneuroblastoma, in the first year of life.

Aside from high-dose radiation and prior chemotherapy, there are no established external risk factors for childhood cancer[3]. There is a growing body of evidence that some harmful in utero or prenatal environmental exposures are associated with increased risk of disease in offspring [4-7]. Air pollution (traffic related air toxics as well as indoor pollution) has been strongly associated with pediatric cancers[8,9]. Fine particle air pollution (PM2.5), a component of outdoor and indoor air pollution originating from both urban and rural sources[10], has carcinogenic properties[11-13]. A comprehensive analysis of adult and pediatric incidence cancer cases of 21 different cancers showed an overall positive association with PM2.5 and cancer, particularly among oral, rectal, lung, skin, and kidney cancers [14]. In addition to air pollution, pesticides are a second environmental exposure relevant to childhood cancer risk. Like air pollution, pesticides are a ubiquitous exposure given their widespread and frequent use in homes and agricultural settings[15]. Pesticides are typically referred to in research by functional class (herbicides, fungicides or insecticides), though sometimes are grouped by compounds (organophosphate pesticides, triazine herbicides, etc.)[16]. Pesticides are known to pose a risk to human health due to their neurotoxic[17] and endocrine disruptive properties[18]. Evidence of pesticide exposure and cancer risk is robust with insecticide exposures during early life appearing to be the most significant with the highest risk of exposure observed during pregnancy[7]. Evidence of a dose-response effect among children exposed to insecticides reinforces their potentially high penetrance of the in-utero environment, potentially leading to early childhood cancers[7].

Social factors and behaviors might also play a role in the development of embryonal cancer, though the social patterning and etiology largely unknown[19]. Embryonal cancer incidence is heterogeneous across race and economic class in the United States[3]. A census-tract level of examination of embryonal tumors using the Surveillance, Epidemiology and End Results (SEER) database found that the incidence of embryonal cancer differed by socioeconomic status (SES) after controlling for race and ethnicity[19]. For all embryonal cancers, white children have the highest incidence of disease, but subgroups of embryonal cancers show differing incidence ratios[19,20]. For example, retinoblastoma is more common among Hispanics and neuroblastoma is most common among white children[20]. Significant, increasing trends have been observed in cancer incidence by SES for rhabdomyosarcoma and hepatoblastoma[19]. Smoking, a social and behavioral factor, may play a role in the development of embryonal cancers in young children. Chemicals in tobacco smoke are known to be carcinogenic[21], cause DNA damage[22] and cross the placenta during pregnancy[23]. Maternal substance use, including smoking, is associated with a marginally increased cancer incidence rate compared to children unexposed to maternal substance use[24]. A population-based cohort study found significantly increased odds of cancer among children whose fathers smoked at the time of conception[25]. The International Agency for Research on Cancer (IARC) has cited obesity as a risk factor for several cancer types[26].

Children born to mothers with high body mass index (BMI) have been found to be associated with higher risk of pediatric cancer and risk of pediatric neuroblastoma was associated with mothers who experienced weight gain during pregnancy[27]. These results suggest maternal obesity may play a significant example of early-life exposures related to childhood cancer development.

Given the rarity and young age of embryonic cancer patients, investigating risk factors and establishing associations for embryonic cancer is particularly difficult. However, examining environmental and social risk factors may elucidate childhood cancer risk. Building evidence for modifiable risk factors is needed if we are to make efforts to reduce harmful exposures and consequently reduce cancer risk.

This ecological study uses data of 875 embryonal cancer cases registered in the state of Georgia from 2003-2017 to examine the joint effects of air pollution, pesticide, and social risk factors on embryonal cancer incidence among 0-4-year-olds.

Methods

Georgia Cancer Registry

The geospatial clustering of pediatric cancer cases in Georgia, USA, was investigated using geocoded incidences of pediatric cancer from the Georgia cancer registry. Incidence rates of patients included those aged 0 to 4 years who had a new diagnosis of one of the seven subgroups of embryonal cancer between 2003 and 2017: IIIc Intracranial and Intraspinal Embryonal Cancers, Iva Neuroblastoma and ganglioneuroblastoma, V Retinoblastoma, Via Nephroblastoma (Wilms Tumor) and other non-epithelial renal tumors, VIIa: Hepatoblastoma and mesenchymal tumors of the liver, IXa Rhabdomyosarcomas, XIIa: Other specified malignant tumors (e.g. pulmonary and pleuropulmonary blastoma) grouped according to the International Classification of Childhood Cancer (ICCC) based on ICD-O-3/WHO 2008[28]. Due to very low case numbers in the subgroups of embryonal cancer, we aggregated all subgroups to analyze the standardized incident ratio of Embryonal Cancers as a whole.

Air Pollution Exposures

High-quality estimates for $PM_{2.5}$ and its components (black carbon (BC), sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), organic matter (OM), mineral dust (DUST), and sea-salt (SS)) with a 1 km resolution were used, which are publicly available[29]. The estimates were derived from a chemical transport model (GEOSChem) and satellite observations of aerosol optical depth, and statistically fused with ground-based observations using a geographically weighted regression over North America to produce a spatially complete representation of $PM_{2.5}$ components. Population-weighted averages for each county were calculated for the years 2000-2016.

In addition, we used previously estimated daily ambient ozone (O_3) concentrations from 2000 to 2016[30]. Exposure concentrations were estimated at 1 km spatial resolution using a well-validated ensemble machine learning model, which integrated multiple predictor variables and three machine learning algorithms[30]. Daily concentrations in each county were estimated based

on the annual predictions at 1 km^2 grid cells by calculating population-weighted averages of these gridded predictions whose centroids fall within the boundary of a given county.

Pesticides

Country-level data on pesticides were collected from the USDA National Agricultural Statistics Service searchable database called Quick Stats. For the years 1997-2017, each type of pesticide (chemical fungicide, herbicide, chemical insecticide (excluding nematicide), chemical insecticide (nematicide), other chemicals) had self-reported number of acres treated. The number of acres treated with pesticides were divided by the total number of acres in the county to create a 'percentage of county treated' variable. Pesticide categories with more than 10% missing country data were excluded from models. Categories with less than 10% were imputed by a simple average of the surrounding counties. Information on the pesticide categories chemical fungicide, herbicide, and chemical insecticide (excluding nematicide) were included in the final analysis.

Neighborhood socioeconomic deprivation

Socioeconomic characteristics, also known as neighborhood socioeconomic deprivation, was measured using 17 county-level indicators corresponding to six socioeconomic domains [poverty/income, racial composition, education, employment, occupation, and housing properties]. These indicators include information on percent males and females with less than a high school education, unemployment rate, percent males no longer in household, percent vacant households, percent households with more than one occupant per room, median household value, percent males and females not in management/business/science/art occupation, percent in poverty, percent female-headed with dependent children, percent household earning under $35,000 per year, percent households on public assistance, percent household with no car, percent households that are non-Hispanic black, and percent households who have the same residence since 2010). Data were accessed through the R-package *tidycensus*, which pulls data from the U.S. Census Bureau and the American Community Survey.

Behavioral factors

Annual county-level estimates for cigarette smoking based on data on 4.7 million adults age 18 and older was obtained from the Behavioral Risk Factor Surveillance System (BRFSS) of 1996 to 2012 (Dwyer-Lindgren et al., 2014). Estimated prevalence of drinking patterns for each US county between 2002 and 2012 and estimated obesity and physical activity prevalence for 2001 to 2011 are also available from the BRFSS (Dwyer-Lindgren et al., 2015, 2013).

Statistical analysis

We calculated site-specific standardized incidence ratios (SIR) of embryonal cancers for each county during the study period (2003-2017) to ensure that we had enough cases for the analyses. The SIR is calculated by dividing the observed number of cases by the expected number of cases in each county. Expected cases per year are estimated by multiplying the national incidence rates for each race-sex subgroup (obtained from Surveillance, Epidemiology, and End Results (SEER) Program) by the corresponding number of individuals in each subgroup

living in each county[31] then aggregating the expected cases for all subgroups. Numbers of individuals within each subgroup (race and sex) were obtained from the US Census Data.

It is possible joint effects of different environmental and social factors impact the differences in incident rates between counties, so we applied an unsupervised learning tool known as self-organizing maps (SOM) to identify pertinent types of multipollutant combinations[32,33]. SOM is unique as it simultaneously applies optimized clustering and data projection to identify data-driven profiles (i.e., categories, clusters) and 'map' them in spatially organized array that results in similarity among neighbors. The benefits of the SOM include the 'map' for visualization, which supports a better understanding of the relationships between the profiles – a huge benefit to air pollution mixture studies. The number of categories chosen for the SOM is important (like other clustering techniques) and determined by the user[34].

The goal was to balance the number of multipollutant categories that minimized the information lost in our exposure of interest with limiting the number of clusters distinct enough to facilitate interpretation of differences between clusters. The number of clusters identified by the SOM algorithm was determined by identifying group structure using within cluster sum of squares and between cluster sum of squares statistics, as well as visual inspection of the cluster star plot. These methods identify clusters with exposure levels homogenous within the cluster and heterogenous between clusters.

Finally, Leroux CAR models are fit with indicators for each multipollutant category as exposure variable and the SIR of embryonal cancers as outcome. The reference category was defined as the multipollutant category with the lowest levels of adverse factors such as low air pollution, pesticide concentrations and high SES. The result is a model describing the joint effect of observed mixture combinations of environmental, socioeconomic, and behavioral factors on spatial clustering of embryonal cancer.

As embryonal cancers are rare, we checked the SIR for zero-inflation using histograms. In case of zero-inflation, we will run Leroux CAR models with zero-inflated Poisson distribution in addition to the traditional Poisson distribution.

All analyses were done using R 4.1.0.

Results

Description of study population

Our analysis was based on data from 875 patients aged 0-4 years old registered in the Georgia Cancer Registry from 2003-2017 who had a new diagnosis of one of the seven subgroups of embryonal cancer (Table 1). The most common subtypes of embryonal cancer were neuroblastoma and ganglioneuroblastoma (279 cases [31.9%]), nephroblastoma and other non-epithelial renal tumors (205 [23.4%]). 151 cases (17.3%) were diagnosed with retinoblastoma and 108 (12.3%) with intracranial and intraspinal embryonal tumors. Three subgroups represented less than 10% of the data: hepatoblastoma and mesenchymal tumors of the liver (52 [5.9%]), rhabdomyosarcomas (73 [8.5%]), and other specified malignant tumors (e.g.,

pulmonary and pleuropulmonary blastoma). Most of the cases were of white ancestry (453 [52%]), followed by those of Black ancestry (300 [34%]). The average age (Standard Deviation) at diagnosis of the cases was 1.6 years (SD 1.3 years) and about half of the cases were male (453 [52%]).

Figure 1 shows the geographic distribution of embryonal cases and SIRs in the state of Georgia. Both the case count and SIRs of embryonal cancers followed a right skewed distribution. The median number of cases per county was 2.0 with an IQR of 4.5 and the median SIR per county was 0.92 with an IQR of 1.13 (Table 2). When observed cases were standardized by expected cases given age, sex and race distribution of each county using the SIR, the incidence of embryonal cancer was higher than expected in rural suburban areas (Figure 1B). Out of 159 counties in Georgia, 40 (25%) counties observed zero cases of embryonal cancer in the 2003–2017-year period (Figure 1C). A description of all air pollutants, pesticide groups, socioeconomic, and behavioral factors is detailed in Table 2.

Self-Organizing Maps (SOM)

5 clusters of environmental (air pollution and pesticides) socioeconomic and behavioral factors in the counties were identified by the SOM (Figure 2A). Each cluster is composed of the median value of a contributing component and an entire cluster. Rural counties are represented by SOM Cluster 1-3. SOM Cluster 4 is composed of 32 suburban counties and SOM Cluster 5 includes four urban counties in the Atlanta Metro area (Figure 2B).

Median (IQR) $PM_{2.5}$ concentrations were highest in urban counties (SOM Cluster 5, 11.1(0.82)) and lowest in rural counties (SOM Cluster 1, 9.42(0.29)) (Table 2). Concentrations of fungicide, herbicide and pesticide were all higher in rural areas (SOM Cluster 1 and SOM Cluster 2). Lower exposures of fungicide, herbicide and pesticide use were observed in urban counties (SOM Cluster 5). For example, SOM Cluster 1 had a median percent land treated with herbicide of 9.79(11.83) and SOM Cluster 5 of 1.64(2.55), while median percent land treated with insecticide in SOM Cluster 1 was 18.36(11.47) and SOM Cluster 5 saw a median percent land treatment of 0.28(0.39). Rural counties showed indications of lower SES compared with urban and suburban counties. SOM Cluster 1-3 showed higher proportions of the population living in poverty and more people with education less than high school. Alcohol consumption was higher overall in urban areas, whereas smoking patterns were more prevalent in rural counties. Higher proportion of obesity was also observed in rural counties.

SOM Cluster 4 was used as the reference for the Leroux CAR models using indicators of SOM cluster membership as exposure because it had the lowest level of adverse risk factors (Figure 2A). Leroux CAR models with zero-inflated Poisson distribution were used for associations with embryonal cancer because the distribution of the SIR of embryonal cancer showed zero-inflation (Figure 1C+D). SOM Cluster 3 (0.73 [0.52, 0.99]) showed a significantly lower SIR than in SOM Cluster 4. Proportions of alcohol consumption (across type and gender) were lower in SOM Cluster 3 compared to SOM cluster 4. Other factors (environmental, SES and behavioral) were at least as common in SOM Cluster 3 as in cluster 4. There was no significant difference in SIRs of embryonal cancer between SOM Clusters 1, 2, or 5 and SOM Cluster 4 (reference).

Discussion

In this ecological study of 159 counties in Georgia, United States, data from 875 pediatric embryonal cancer cases registered in Georgia Cancer Registry from 2003 to 2017 were aggregated for analysis. Application of the spatial clustering algorithm SOM was used to assess the association between environmental and social/behavioral risk factor. Significantly higher SIRs were observed in areas of increased alcohol consumption. To our knowledge, this is one of the first studies to find a significant association between an alcohol consumption and embryonal cancer.

Rarity is a central fact of any study examining embryonal cancers. As such, it is difficult to build causal associations or assess risk factors when low numbers dictate the quality and quantity of evidence[3]. Case control study designs, while susceptible to recall and selection biases[35], have been the bulk of etiologic investigations of embryonal cancers. Case control designs are not effective at environmental exposure assessments, especially those like air pollution and pesticide exposures, which are nearly ubiquitous in our environment[36]. While they can collect individual characteristics, such as demographics, via questionnaire, they are inadequate assessments of wider environmental contexts. Due to the early onset of embryonal cancer, parental health behaviors are a key to the investigation of cancers in very young children[25]. While this study does not include information on parental health behaviors of embryonal cancer cases, it leverages county-level averages as an approximation of the prenatal environment. Using county-level data allowed us to aggregate data from all cancer cases that were diagnosed in each county in a specific time window. This does not only increase the number of cases to be included in the analysis but also avoids any potential selection bias.

Alcohol has been established as a risk factor for several cancers in adults[37], however evidence for parental alcohol consumption and pediatric cancer is inconsistent[38-40]. Since byproducts and metabolites of alcohol are known or suspected carcinogens[41], and are known to cross the placental barrier in pregnant women[42], investigations of this exposure are crucial. Other suggested mechanisms for the apparent carcinogenic effect of alcohol include facilitation of cellular entry for other carcinogens, induction of carcinogen-activating enzymes, and inhibition of DNA synthesis and repair[43-45]. Parental alcohol consumption has been consistently associated with significant, increased risk of some pediatric cancers[46], but to date subgroups of embryonal cancers have not been consistently associated with prenatal alcohol exopsure[39,46]. A population-based cohort study examining prenatal substance use included several types of embryonal cancer (central nervous system tumors, retinoblastoma, rhabdomyosarcomas, and neuroblastoma), but did not find an association between alcohol consumption and risk of these cancers- though they noted imprecise estimates due to low case numbers[24]. Alcohol was associated to the spatial clustering of pediatric lymphoma in Georgia using similar methodology, however, embryonal cancer was not included in that analysis[47].

Prior studies have shown associations between air pollution exposure and embryonal cancer[4,6], however results of this study did not yield significant associations with air pollution components. Air pollution exposure in the United States is heterogeneous across race and income groups. While absolute concentrations of air pollution declined from 1990-2010, racial/ethnic disparities remained across income levels, in urban and rural areas, and in all states, for multiple pollutants at the end of that time period[48]. Fine particle air pollution

(PM2.5), a component of outdoor and indoor air pollution originating from both urban and rural sources[10], has carcinogenic properties[11-13] due to its biological mechanisms of systemic inflammation[11] and DNA damage[13]. While this study did not show a relationship with air pollution, it is likely that county-level data for air pollution does not provide a sufficient spatial resolution to find associations. While census-tract level estimates of exposures (or at an even finer resolution) may, in general, offer more accurate measures of air toxics, embryonal cancer incidence would be too small to derive any type of association.

This study has several limitations. Our ecological study design used county level estimates of exposure concentration and case counts of embryonal cancer. Individual level data was only used for descriptive analysis of the children's race, sex, and age at the time of diagnosis. Low case counts of embryonal cancer subtypes necessitated their aggregation to a single group for the main analysis, despite the fact the fact that incidence varied between subtype, sex, and race. The exposure modeling was done using the residential address of the children at time of diagnosis; however, this does not necessarily reflect critical periods of exposure. However, there is evidence that family mobility does not result in a significant change in neighborhood[51,52] or in neighborhood environmental exposure assessment[53]. Furthermore, due to the joint effect study design of the SOM, single effects could not be calculated. While this clustering method is useful for increasing statistical power, it cannot estimate the effects of individual exposures. Finally, this study was limited to the state of Georgia and those children diagnosed with embryonal cancer between the ages of 0 and 4 years old. This limits the generalizability of the study to other cancer groups, age ranges and locations. Further studies should be conducted on a regional or nationwide level to increase the cases and the spatial variation of exposures.

In conclusion, this etiologic study focused on the spatial clustering of external risk factors associated with embryonal cancer. Our findings suggest an association between alcohol and the SIR of embryonal cancer in Georgia. Future studies should be conducted to replicate these findings in wider geographic areas and using individual-level data.

References

1. (CDC) CfDCaP. Ten Leading Causes of Death and Injury. National Center for Injury Prevention and Control 2020. Accessed March 23, 2022, 2022. https://wisqars.cdc.gov/cgi-bin/broker.exe
2. Tulla M, Berthold F, Graf N, et al. Incidence, Trends, and Survival of Children With Embryonal Tumors. *Pediatrics*. 2015;136(3):e623-e632. doi:10.1542/peds.2015-0224
3. Spector LG, Pankratz N, Marcotte EL. Genetic and nongenetic risk factors for childhood cancer. *Pediatr Clin North Am*. Feb 2015;62(1):11-25. doi:10.1016/j.pcl.2014.09.013
4. Heck JE, Park AS, Qiu J, Cockburn M, Ritz B. Retinoblastoma and ambient exposure to air toxics in the perinatal period. *J Expo Sci Environ Epidemiol*. Mar-Apr 2015;25(2):182-6. doi:10.1038/jes.2013.84
5. Na Zhu XJ, Xilin Geng, Huifeng Yue, Guangke Li, Nan Sang. Maternal PM2.5 exposure and abnormal placental nutrient transport. *Ecotoxicology and Environmental Safety*. 2021;207doi:https://doi.org/10.1016/j.ecoenv.2020.111281
6. Ghosh JK, Heck JE, Cockburn M, Su J, Jerrett M, Ritz B. Prenatal exposure to traffic-related air pollution and risk of early childhood cancers. *Am J Epidemiol*. Oct 15 2013;178(8):1233-9. doi:10.1093/aje/kwt129
7. Ma X, Buffler PA, Gunier RB, et al. Critical windows of exposure to household pesticides and risk of childhood leukemia. *Environ Health Perspect*. 2002;110(9):955-960. doi:10.1289/ehp.02110955
8. Hall C, Heck JE, Ritz B, Cockburn M, Escobedo LA, von Ehrenstein OS. Prenatal Exposure to Air Toxics and Malignant Germ Cell Tumors in Young Children. *J Occup Environ Med*. Jun 2019;61(6):529-534. doi:10.1097/JOM.0000000000001609
9. Filippini T, Heck JE, Malagoli C, Del Giovane C, Vinceti M. A review and meta-analysis of outdoor air pollution and risk of childhood leukemia. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev*. 2015;33(1):36-66. doi:10.1080/10590501.2015.1002999
10. Tucker WG. An overview of PM2.5 sources and control strategies. *Fuel Processing Technology*. 2000/06/01/ 2000;65-66:379-392. doi:https://doi.org/10.1016/S0378-3820(99)00105-8
11. Loomis D, Grosse Y, Lauby-Secretan B, et al. The carcinogenicity of outdoor air pollution. *The Lancet Oncology*. 2013;14(13):1262-1263. doi:10.1016/S1470-2045(13)70487-X
12. Risom L, Møller P, Loft S. Oxidative stress-induced DNA damage by particulate air pollution. *Mutat Res*. Dec 30 2005;592(1-2):119-37. doi:10.1016/j.mrfmmm.2005.06.012
13. Newby DE, Mannucci PM, Tell GS, et al. Expert position paper on air pollution and cardiovascular disease. *European Heart Journal*. 2014;36(2):83-93. doi:10.1093/eurheartj/ehu458
14. Coleman NC, Burnett RT, Ezzati M, Marshall JD, Robinson AL, Pope CA, 3rd. Fine Particulate Matter Exposure and Cancer Incidence: Analysis of SEER Cancer Registry Data from 1992-2016. *Environ Health Perspect*. Oct 2020;128(10):107004. doi:10.1289/EHP7246
15. Zhao F, Li L, Chen Y, et al. Risk-Based Chemical Ranking and Generating a Prioritized Human Exposome Database. *Environ Health Perspect*. Apr 2021;129(4):47014. doi:10.1289/EHP7722

16. Alavanja MCR, Hoppin JA, Kamel F. Health Effects of Chronic Pesticide Exposure: Cancer and Neurotoxicity. *Annual Review of Public Health*. 2004;25(1):155-197. doi:10.1146/annurev.publhealth.25.101802.123020
17. Keifer MC, Mahurin RK. Chronic neurologic effects of pesticide overexposure. *Occup Med*. Apr-Jun 1997;12(2):291-304.
18. Tabb MM, Blumberg B. New modes of action for endocrine-disrupting chemicals. *Mol Endocrinol*. Mar 2006;20(3):475-82. doi:10.1210/me.2004-0513
19. Geris JM, Spector LG. Race, ethnicity, and socioeconomic differences in incidence of pediatric embryonal tumors in the United States. *Pediatr Blood Cancer*. Sep 2020;67(9):e28582. doi:10.1002/pbc.28582
20. Friedrich P, Itriago E, Rodriguez-Galindo C, Ribeiro K. Racial and Ethnic Disparities in the Incidence of Pediatric Extracranial Embryonal Tumors. *JNCI: Journal of the National Cancer Institute*. 2017;109(10)doi:10.1093/jnci/djx050
21. IARC. Tobacco smoke and involuntary smoking. *IARC Monogr Eval Carcinog Risks Hum*. 2004;83:1-1438.
22. Fraga CG, Motchnik PA, Wyrobek AJ, Rempel DM, Ames BN. Smoking and low antioxidant levels increase oxidative damage to sperm DNA. *Mutat Res*. Apr 13 1996;351(2):199-203. doi:10.1016/0027-5107(95)00251-0
23. de la Chica RA, Ribas I, Giraldo J, Egozcue J, Fuster C. Chromosomal Instability in Amniocytes From Fetuses of Mothers Who Smoke. *JAMA*. 2005;293(10):1212-1222. doi:10.1001/jama.293.10.1212
24. Auger N, Goudie C, Low N, Healy-Profitos J, Lo E, Luu TM. Maternal use of illicit drugs, tobacco or alcohol and the risk of childhood cancer before 6 years of age. *Drug Alcohol Depend*. Jul 1 2019;200:133-138. doi:10.1016/j.drugalcdep.2019.03.008
25. Milne E, Greenop KR, Scott RJ, et al. Parental Prenatal Smoking and Risk of Childhood Acute Lymphoblastic Leukemia. *American Journal of Epidemiology*. 2011;175(1):43-53. doi:10.1093/aje/kwr275
26. Vainio H, Kaaks R, Bianchini F. Weight control and physical activity in cancer prevention: international evaluation of the evidence. *Eur J Cancer Prev*. Aug 2002;11 Suppl 2:S94-100.
27. Stacy SL, Buchanich JM, Ma Z-Q, et al. Maternal Obesity, Birth Size, and Risk of Childhood Cancer Development. *American journal of epidemiology*. 2019;188(8):1503-1511. doi:10.1093/aje/kwz118
28. Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, third edition. *Cancer*. Apr 1 2005;103(7):1457-67. doi:10.1002/cncr.20910
29. van Donkelaar A, Martin RV, Li C, Burnett RT. Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors. *Environ Sci Technol*. Mar 5 2019;53(5):2595-2611. doi:10.1021/acs.est.8b06392
30. Requia WJ, Di Q, Silvern R, et al. An Ensemble Learning Approach for Estimating High Spatiotemporal Resolution of Ground-Level Ozone in the Contiguous United States. *Environmental Science & Technology*. 2020/09/15 2020;54(18):11037-11047. doi:10.1021/acs.est.0c01791

31. Clough L, Bayakly AR, Ward KC, et al. Clustering of cutaneous T-cell lymphoma is associated with increased levels of the environmental toxins benzene and trichloroethylene in the state of Georgia. *Cancer*. 2020;126(8):1700-1707. doi:https://doi.org/10.1002/cncr.32665
32. Pearce JL, Waller LA, Mulholland JA, et al. Exploring associations between multipollutant day types and asthma morbidity: epidemiologic applications of self-organizing map ambient air quality classifications. *Environ Health*. Jun 23 2015;14:55. doi:10.1186/s12940-015-0041-8
33. Pearce JL, Waller LA, Sarnat SE, et al. Characterizing the spatial distribution of multiple pollutants and populations at risk in Atlanta, Georgia. *Spat Spatiotemporal Epidemiol*. Aug 2016;18:13-23. doi:10.1016/j.sste.2016.02.002
34. Pearce JL, Waller LA, Chang HH, et al. Using self-organizing maps to develop ambient air quality classifications: a time series example. *Environmental Health*. 2014/07/03 2014;13(1):56. doi:10.1186/1476-069X-13-56
35. Wacholder S. Design issues in case-control studies. *Stat Methods Med Res*. Dec 1995;4(4):293-309. doi:10.1177/096228029500400403
36. Wambaugh JF, Wang A, Dionisio KL, et al. High Throughput Heuristics for Prioritizing Human Exposure to Environmental Chemicals. *Environmental Science & Technology*. 2014/11/04 2014;48(21):12760-12767. doi:10.1021/es503583j
37. Raquel Ferreira de Menezes AB, Luiz Claudio Santos Thuler. Alcohol Consumption and Risk of Cancer: a Systematic Literature Review. *Asian Pacific Journal of Cancer Prevention*. 2013;14(9):4965-4972.
38. Orsi L, Rudant J, Ajrouche R, et al. Parental smoking, maternal alcohol, coffee and tea consumption during pregnancy, and childhood acute leukemia: the ESTELLE study. *Cancer Causes Control*. Jul 2015;26(7):1003-17. doi:10.1007/s10552-015-0593-5
39. Yang Q, Olshan AF, Bondy ML, et al. Parental smoking and alcohol consumption and risk of neuroblastoma. *Cancer Epidemiol Biomarkers Prev*. Sep 2000;9(9):967-72.
40. Johnson KJ, Williams KS, Ross JA, et al. Parental tobacco and alcohol use and risk of hepatoblastoma in offspring: a report from the children's oncology group. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2013;22(10):1837-1843. doi:10.1158/1055-9965.EPI-13-0432
41. Cancer IAfRo. Alcohol drinking. *IARC monographs on the evaluation of carcinogenic risks to humans*. 1988;44
42. Baan R, Straif K, Grosse Y, et al. Carcinogenicity of alcoholic beverages. *The Lancet Oncology*. 2007;8(4):292-293. doi:10.1016/S1470-2045(07)70099-2
43. Anderson LM, Chhabra SK, Nerurkar PV, Souliotis VL, Kyrtopoulos SA. Alcohol-related cancer risk: A toxico kinetic hypothesis. *Alcohol*. 1995/03/01/ 1995;12(2):97-104. doi:https://doi.org/10.1016/0741-8329(94)00089-1
44. Lieber CS, Seitz HK, Garro AJ, Worner TM. Alcohol-related diseases and carcinogenesis. *Cancer research*. 1979;39(7 Part 2):2863-2886.
45. Mufti SI, Darban HR, Watson RR. Alcohol, cancer, and immunomodulation. *Critical reviews in oncology/hematology*. 1989;9(3):243-261.
46. Infante-Rivard C, El-Zein M. Parental Alcohol Consumption and Childhood Cancers: A Review. *Journal of Toxicology and Environmental Health, Part B*. 2007/01/01 2007;10(1-2):101-129. doi:10.1080/10937400601034597

47. Anke Huels YL, Grace Christensen, Zhenjiang Li, John L. Pearce, Rana Bayakly, Timothy L. Lash, Kevin Ward, and Jeffrey M. Switchenko. Environmental and social risk factors in association with spatial clustering of childhood cancer incidence. *ISEE Conference Abstracts*. 2021;2021(1)
48. Liu J, Clark LP, Bechle MJ, et al. Disparities in Air Pollution Exposure in the United States by Race/Ethnicity and Income, 1990–2010. *Environ Health Perspect*. 2021;129(12):127005. doi:doi:10.1289/EHP8584
49. P. B. *CALINE4: A Dispersion Model for Predicting Air Pollutant Concentrations Near Roadways*. 1984. 1989. https://trid.trb.org/view/215944
50. Danysh HE, Zhang K, Mitchell LE, Scheurer ME, Lupo PJ. Maternal residential proximity to major roadways at delivery and childhood central nervous system tumors. *Environ Res*. Apr 2016;146:315-22. doi:10.1016/j.envres.2016.01.012
51. Chen L, Bell EM, Caton AR, Druschel CM, Lin S. Residential mobility during pregnancy and the potential for ambient air pollution exposure misclassification. *Environmental Research*. 2010/02/01/ 2010;110(2):162-168. doi:https://doi.org/10.1016/j.envres.2009.11.001
52. Lupo PJ, Symanski E, Chan W, et al. Differences in exposure assignment between conception and delivery: the impact of maternal mobility. *Paediatric and Perinatal Epidemiology*. 2010;24(2):200-208. doi:https://doi.org/10.1111/j.1365-3016.2010.01096.x
53. Danysh HE, Mitchell LE, Zhang K, Scheurer ME, Lupo PJ. Differences in environmental exposure assignment due to residential mobility among children with a central nervous system tumor: Texas, 1995-2009. *J Expo Sci Environ Epidemiol*. Jan 2017;27(1):41-46. doi:10.1038/jes.2015.63

Tables and Figures

Table 1: Distribution of embryonal cancer cases registered in Georgia Cancer Registry from 2003 to 2017, stratified by subtype and count, age, sex and race.

	Total	IVa[2]	VIa[4]	V[3]	IIIC[1]	IXa[6]	VIIa[5]	XIIa[7]
Count	875 (100%)	279 (31.9%)	205 (23.4%)	151 (17.3%)	108 (12.3%)	73 (8.5%)	52 (5.9%)	7 (0.8%)
Age (mean (SD))	1.6 (1.3)	1.3 (1.3)	1.9 (1.4)	1.0 (1.1)	1.7 (1.3)	2.2 (1.3)	1.5 (1.3)	1.9 (0.9)
Sex (% male)	453 (52%)	165 (59%)	95 (46%)	66 (44%)	51 (47%)	43 (59%)	30 (58%)	3 (43%)
Race								
white	537 (61%)	177 (63%)	113 (55%)	85 (56%)	78 (72%)	48 (66%)	34 (65%)	2 (29%)
black	300 (34%)	90 (32%)	85 (41%)	59 (39%)	28 (26%)	23 (32%)	12 (23%)	3 (43%)
other	38 (4%)	12 (4%)	7 (3%)	7 (5%)	2 (2%)	2 (3%)	6 (12%)	2 (29%)

Cancer types grouped according to the International Classification of Childhood Cancer (ICCC).

[1] IVa: Neuroblastoma and ganglioneuroblastoma

[2] Via: Nephroblastoma (Wilms Tumor) and other non-epithelial renal tumors

[3] V: Retinoblastoma

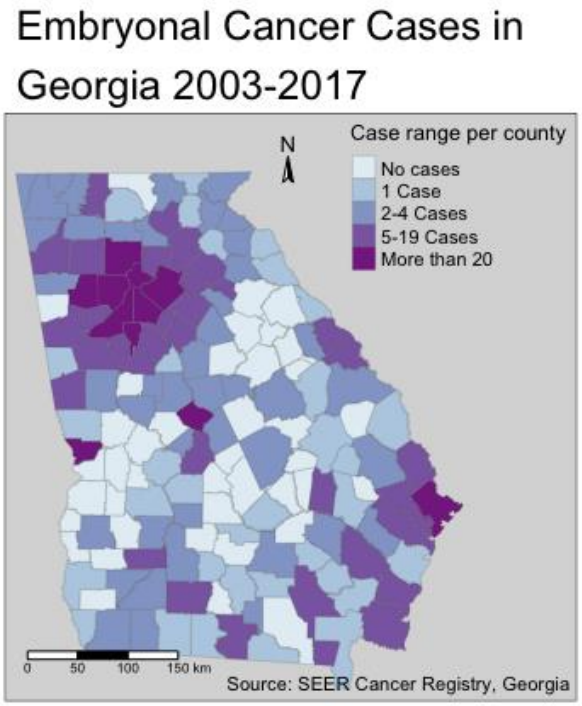
[4] IIIc: Intracranial and Intraspinal Embryonal Cancers

[5] IXa: Rhabdomyosarcomas

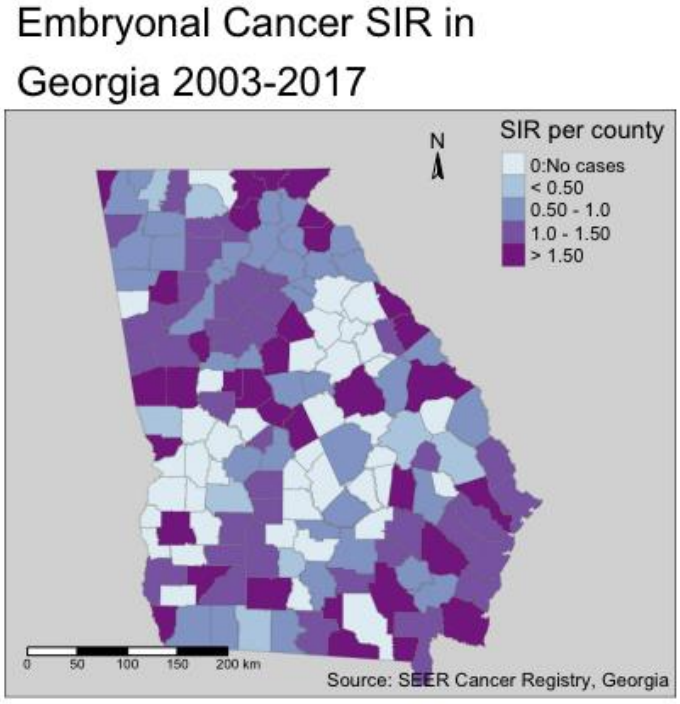
[6] VIIa: Hepatoblastoma and mesenchymal tumors of the liver

[7] XIIa: Other specified malignant tumors (e.g., pulmonary and pleuropulmonary blastoma)

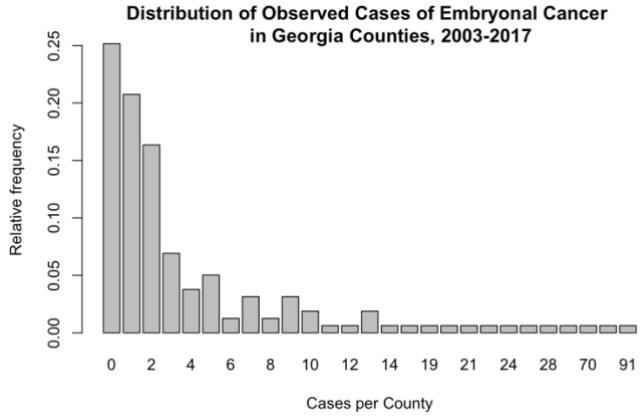
A.



B.



C.



D.

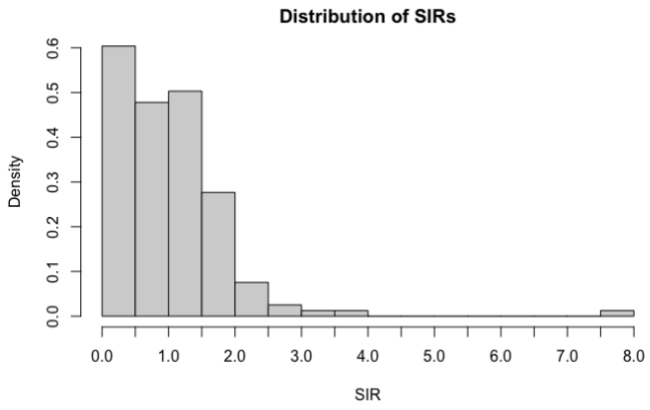
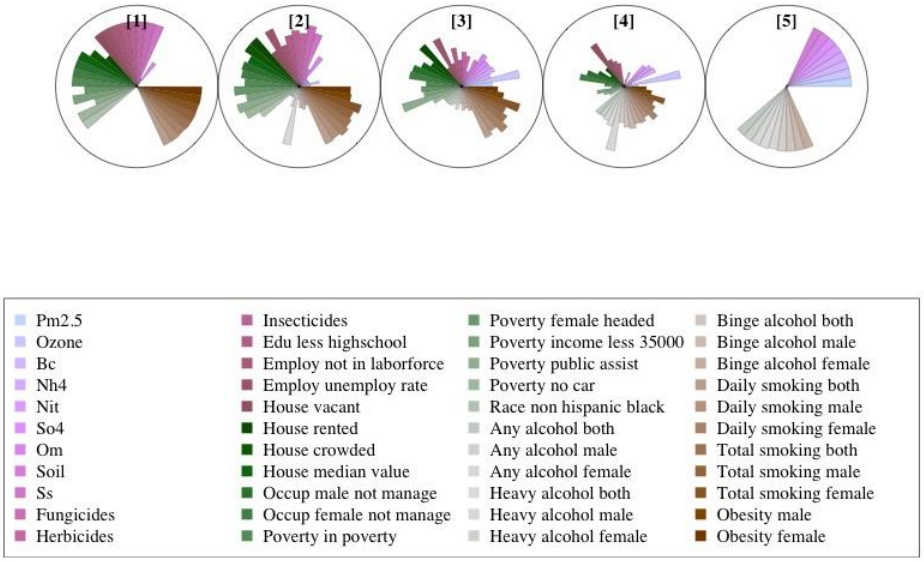


Figure 1. Distribution of embyronal cancer cases aged 0-4 years in Georgia. Panel A: Geographic distriubtion of embryonal cancer cases per county. Panel B: Geographic distribution of the Standardized Incidence Rate (SIR) of embyronal cancer per county. Panel C: Histogram of embryonal cancer cases. Panel D: Histogam of the SIRs of embryonal cancer.

A.



B.

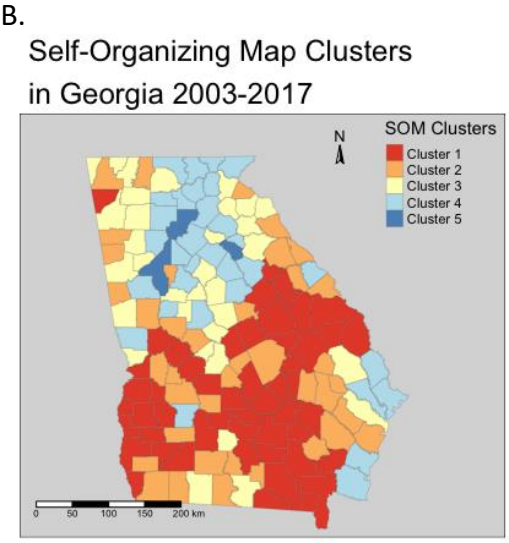


Figure 2. Self-Organizing Maps (SOM). Exposure clusters in association with increased SIR of embryonal cancer. **A.** SOM cluster plot, each circle is a cluster and each piece of the cluster represents the median value of a contributing component. **B**. Geographic distribution of clusters in Georgia. Each color is a cluster.

Table 2. Distribution of study characteristics among the 159 counties in Georgia, United States

	All Counties Median (IQR)	SOM Cluster 1 Median (IQR)	SOM Cluster 2 Median (IQR)	SOM Cluster 3 Median (IQR)	SOM Cluster 4 Median (IQR)	SOM Cluster 5 Median (IQR)
N - Counties	159	51	39	33	32	4
SIRs of embryonal cancer[1]	0.92 (1.13)	0.49(1.06)	0.86(1.01)	0.92(0.89)	1.19(0.68)	0.9(0.32)
$PM_{2.5}$, median (IQR)	9.61 (0.75)	9.42(0.29)	9.58(0.48)	10.04(1.09)	9.85(1.27)	11.1(0.82)
$PM_{2.5}$ components, median (IQR)						
Black carbon (BC)	1 (0.1)	0.98(0.05)	1(0.08)	1.06(0.11)	1.02(0.18)	1.13(0.11)
Ammonium (NH_4)	5.71 (2.33)	4.99(0.62)	5.17(1.44)	7.22(1.75)	6.86(1.86)	8.48(1.11)
Nitrate (NO_3)	3.14 (2.2)	2.8(0.33)	3.01(1)	4.96(2.36)	4.47(2.48)	6.33(1.75)
Organic matter (OM)	4.37 (0.31)	4.37(0.19)	4.37(0.3)	4.43(0.37)	4.29(0.59)	4.57(0.4)
Sulfate (SO_4)	2.45 (0.36)	2.3(0.13)	2.36(0.26)	2.65(0.28)	2.59(0.28)	2.96(0.24)
Mineral dust (DUST)	6.42 (0.66)	6.34(0.4)	6.43(0.54)	6.63(0.8)	6.29(1.03)	7.23(1.08)
Sea-salt (SS)	2.97 (1.32)	3.38(0.68)	3.2(1.26)	2.34(0.94)	2.46(0.74)	2.45(0.31)
Ozone, median (IQR)	40.13 (1.39)	39.72(0.83)	39.95(1.33)	40.83(1.28)	40.86(1.28)	40.72(0.42)

Pesticides, median (IQR)						
Chemical fungicide (% land treated)	0.49 (3.2)	2.12(3.98)	2.15(4.4)	0.07(0.63)	0.04(0.13)	0.05(0.12)
Herbicide (% land treated)	3.27 (9.41)	9.79(11.83)	4.27(11.12)	3.17(6.42)	1.26(1.82)	1.64(2.55)
Chemical insecticide (% land treated)	1.26 (8.66)	8.36(11.47)	3.39(11.46)	0.69(1.47)	0.3(0.42)	0.28(0.39)
Neighborhood SES, median (IQR)						
% Education less than high school	21.62 (7.89)	25.46(5.92)	22.87(4.58)	19.47(5.94)	14.58(5.14)	8.62(0.92)
% Not in labor force	37.39 (12.22)	45.49(9.17)	37.69(4.95)	33.95(6.76)	31.59(11.07)	26.14(2.03)
Unemployment rate	9.53 (3.25)	10.23(4.34)	10.91(3.54)	10.01(3)	8.42(1.54)	6.65(1.65)
% Homes crowded	2.29 (1.31)	2.49(1.46)	2.94(1.09)	2.18(0.79)	1.71(0.54)	1.14(0.28)
Median home value ($ in thousands)	$103,266.67 ($58,155.55)	$70,988.89 ($11,977.78)	$92,088.89 ($16,011.11)	$125,077.78 ($12,266.67)	$168,222.22 ($24,283.34)	$245,288.89 ($14,441.67)
% Homes rented	29.98 (9.34)	30.03(6.33)	32.38(8.35)	31.34(11.91)	25.29(8.82)	18.18(9.3)
% Homes vacant	16.09 (7.18)	19.28(5.71)	15.74(4.92)	13.2(4.87)	12.57(11.67)	8.68(3.54)
% Female not in management	65.88 (5.77)	68.08(5.41)	67.14(4.48)	64.81(3.07)	62.13(6.06)	51.44(1.68)
% Male not in management	77.96 (6.48)	80.99(4.89)	78.8(5.51)	76.08(4.57)	71.85(6.21)	53.27(2.77)
% Female headed households	8.02 (2.76)	8.59(2.62)	8.68(2.89)	7.26(3.79)	6.87(3.03)	5.46(1.38)
% In poverty	16.82 (7.59)	20.24(6.84)	18.82(4.4)	14.41(4.78)	10.42(4.04)	5.72(2.38)
% Income less than $35,000	47.04 (12.23)	53.05(7.19)	48.57(5.63)	42.37(5.7)	31.87(12.99)	19.96(5.91)
% No car	22.7 (6.08)	25.98(4.81)	24.6(3.83)	21.16(3.67)	18.95(4.34)	14.93(4.71)
% On public assistance	1.75 (0.84)	1.66(0.88)	2.12(0.78)	1.9(0.7)	1.46(0.59)	1(0.5)
% Non-Hispanic Black	27.93 (25.06)	36.43(23.73)	28.58(15.64)	21.36(24.83)	16.32(20.01)	12.57(21.09)
% moved in before 2010	77.35 (6.76)	80.06(4.28)	76.65(5.97)	74.51(6.65)	75.07(7.43)	76.52(7.16)
Drinking patterns %, median (IQR)						
Any alcohol (both gender)	38.51 (10.08)	34.35(4.44)	37.77(3.84)	41.96(6.9)	47.25(7.94)	58.35(2.17)
Any alcohol (female)	29.64 (10.32)	25.45(3.86)	28.63(3.72)	33.25(6.97)	39.23(7.8)	50.21(2.65)
Any alcohol (male)	47.54 (9.29)	43.07(5.14)	47.1(4.74)	50.59(6.98)	55.55(5.84)	66.81(1.7)

Binge drinking[2] (both gender)	12.78 (2.1)	11.89(1.42)	12.48(1.34)	13.15(1.99)	13.99(1.9)	15.3(1.64)
Binge drinking[2] (female)	7.1 (1.68)	6.1(1.11)	6.99(0.83)	7.59(1.28)	8.28(1.84)	9.78(1.78)
Binge drinking[2] (male)	18.66 (2.42)	17.75(1.89)	18.46(1.96)	18.66(2.12)	19.92(1.86)	21(1.48)
Heavy drinking[3] (both gender)	5.68 (1.23)	5.14(0.92)	5.65(0.93)	5.71(0.54)	6.46(1.2)	6.88(0.76)
Heavy drinking[3] (female)	3.08 (1.24)	2.54(0.61)	3.04(0.61)	3.19(0.94)	4.17(1.31)	5.56(1.14)
Heavy drinking[3] (male)	8.07 (1.45)	7.74(1.22)	8.2(1.18)	7.97(1)	8.56(1.87)	8.3(0.4)
Smoking behavior %, median (IQR)						
Current or ever (both gender)	27.28 (2.83)	28.02(1.6)	27.97(1.46)	26.84(2.13)	25.03(2.66)	18.76(0.58)
Current or ever (female)	23.75 (2.98)	24.02(2.43)	24.97(2.05)	23.81(2.89)	21.88(3.04)	16.59(0.35)
Current or ever (male)	30.86 (3.51)	32.34(1.62)	31.38(1.58)	29.55(1.65)	27.04(1.97)	20.81(1.76)
Current daily (both gender)	21.34 (2.53)	21.85(1.8)	21.93(1.35)	21(2.9)	19.35(2.59)	13.38(0.48)
Current daily (female)	18.91 (2.98)	18.97(2.53)	19.7(2.03)	19.01(3.16)	17.09(3.19)	12.35(0.54)
Current daily (male)	23.91 (3)	24.83(1.29)	24.43(1.39)	23.14(2.68)	21.33(2.13)	14.99(0.92)
Percent obese[4], median (IQR)						
Female	40.94 (7.54)	45.82(5.57)	42.12(3.23)	37.45(4.91)	33.85(4.48)	27.95(1.79)
Male	35.65 (3.62)	37.51(2.34)	35.9(1.9)	34.09(2.45)	31.78(2.66)	27.55(1.37)

Abbreviations: IQR, interquartile range; $PM_{2.5}$, fine particle air pollution; SES, socioeconomic status

[1] Cancer types grouped according to the International Classification of Childhood Cancer (ICCC): IVa: Neuroblastoma and ganglioneuroblastoma, Via: Nephroblastoma (Wilms Tumor) and other non-epithelial renal tumors, V: Retinoblastoma, IIIc: Intracranial and Intraspinal Embryonal Cancers, IXa: Rhabdomyosarcomas, VIIa: Hepatoblastoma and mesenchymal tumors of the liver, XIIa: Other specified malignant tumors (e.g., pulmonary and pleuropulmonary blastoma)

[2] Consuming at least 4 drinks for women or 5 drinks for men on a single occasion at least once in the past 30 days;

[3] Consuming, on average, more than 1 drink per day for women or 2 drinks per day for men in the past 30 days;

[4] BMI ≥30 kg/m^2

Table 3. Results of the Leroux CAR models for the association between SOM cluster membership and SIR of pediatric cancer.

Cancer	Cluster 1	Cluster 2	Cluster 3	Cluster 5
Embryonal Cancers*	0.93 (0.78, 1.10)	0.82 (0.66, 1.25)	0.73 (0.52, 0.99)	0.86 (0.68, 1.10)

Cancer types grouped according to the International Classification of Childhood Cancer (ICCC): IVa: Neuroblastoma and ganglioneuroblastoma, Via: Nephroblastoma (Wilms Tumor) and other non-epithelial renal tumors, V: Retinoblastoma, IIIc: Intracranial and Intraspinal Embryonal Cancers, IXa: Rhabdomyosarcomas, VIIa: Hepatoblastoma and mesenchymal tumors of the liver, XIIa: Other specified malignant tumors (e.g., pulmonary and pleuropulmonary blastoma)
Cluster 4 was used as the reference group because it had the lowest levels of adverse risk factors (compare Figure 2A).
*As the distribution of the SIR of embryonal cancer showed zero-inflation, we used Leroux CAR models with zero-inflated Poisson distribution for associations with embryonal cancers.