Open Access

Proximity to Superfund Sites and Cancer Incidence in the South Atlantic United States

Sarah M. Rockhill, MPH
Kristina L. Boyd, MS
Amy M. Lavery, PhD
Benjamin A. McKenzie, MS
Geospatial Research, Analysis, and Services Program
Office of Innovation and Analytics
National Center for Environmental Health
Agency for Toxic Substances and Disease Registry
Centers for Disease Control and Prevention

Abstract **Public concern about environmental hazards can require health agencies to investigate patterns of cancer that seem unusual; however, measuring relationships between cancer incidence and sources of pollution can be challenging. We present ecological methods for examining cancer incidence rates and residential proximity to Superfund sites in the South Atlantic region of the U.S. We calculated population-weighted average residential proximity to Superfund sites with contaminants linked to colorectal and lung cancer for subcounty geographic areas. Using multilevel linear regression models, we estimated the relative percent change in expected 2014–2018 subcounty age-adjusted incidence rates of colorectal and lung cancers associated with increasing proximity, adjusted for the prevalence of other risk factors for cancers. We observed a 29.2% relative increase (95% confidence interval [CI] [15.0, 45.2]) in the expected subcounty age-adjusted colorectal cancer incidence rate for each one unit increase in Superfund proximity score, but no statistical association between subcounty age-adjusted lung cancer incidence rates and Superfund proximity (β = -1.9%, 95% CI [-5.3%, 1.6%]). Residential proximity to Superfund sites was associated with increased rates of colorectal cancer incidence but not lung cancer incidence. The methods presented in our analysis can be used and adapted by public health agencies to incorporate evaluation of environmental hazards into cancer surveillance practices.**

Keywords: geospatial methods, cancer surveillance, Superfund program, epidemiology

Introduction

Cancers constitute approximately 20% of disability-adjusted life years (DALYs) caused by diseases linked to environmental causes (Pugh & Zarus, 2012). In 2022, the Centers for Disease Control and Prevention (CDC) and the Agency for Toxic Substances and Disease Registry (ATSDR) updated the *Guidelines for Examining Unusual Patterns of Cancer and Environmental Concerns* (hereafter referred to as 2022 Guidelines) to strengthen investigations of cancers and environmental risk factors (Foster et al., 2022). The 2022 Guidelines encourage U.S. state, tribal, local, and territorial (STLT) health departments to proactively monitor cancer registry data to measure trends, identify patterns in cancer incidence, and collaborate with relevant experts to examine potential environmental hazards when unusual patterns of cancer are detected. Although ad hoc cancer cluster investigations are common, Foster et al. report that fewer than one half of STLT health departments report that they proactively evaluate geographic areas for elevated rates of cancer. Frequently cited barriers to proactive investigation by STLT health departments include insufficient staff time, training, access to geospatial software, and incomplete or sparse data. Thus, reproducible methods for evaluating population-level associations between cancer incidence and common sources of pollution could help STLT health departments to proactively identify areas of concern for further investigation.

Areas contaminated with hazardous waste that pose a risk to human and environmental health—classified as Superfund sites by the U.S. Environmental Protection Agency (U.S. EPA, 2011)—can raise community concerns about cancer risks. Approximately 73 million people lived within 3 mi of a Superfund site in 2019, according to estimates by the U.S. EPA, representing approximately 22% of the U.S. population (Office of Land and Emergency Management, 2020).

Environmental contamination from Superfund sites has raised important questions about what health outcomes might be attributed to toxic chemical exposure in neighborhoods surrounding the sites. Several studies have found associations between residential proximity to a particular Superfund site and specific cancer types in some population groups (Iyer et al., 2016; Press et al., 2016; Zhang et al., 2020). Other studies have found that communities near Superfund sites experience an increased risk of noncancer health outcomes such as congenital birth defects and lower life expectancy (Currie et al., 2011;

https://doi.org/10.70387/001c.147152

Kiaghadi et al., 2021). These findings suggest that residential proximity to Superfund sites broadly can be considered a class of environmental exposure, which can confer risk across a range of health outcomes, including cancer.

We aim to demonstrate how publicly available Superfund site data can be used to augment traditional cancer surveillance activities proposed in the 2022 Guidelines. For this demonstration, we leverage data from the CDC National Environmental Public Health Tracking Network (Tracking Network), a centralized data hub that includes county and local data on a wide range of key and emerging environmental hazards and health outcomes, including some cancers associated with environmental causes (Balluz, 2014). The Tracking Network was created by the CDC Environmental Public Health Tracking Program (Tracking Program) to bridge the gap between environmental data and population health and support efforts to better understand possible links between environmental exposures and disease (Balluz, 2014). The overall objective of our study was to demonstrate replicable methods for analysis of proximity-based measures related to environmental hazards.

Methods

With data provided by the Tracking Program, we analyzed age-adjusted subcounty incidence rates for lung and bronchus cancers (hereafter referred to as lung cancers) and colorectal cancers between the years 2014 and 2018. The Tracking Program created subcounty areas by spatially aggregating U.S. Census tracts within states (excluding areas without any residents) into a contiguous area with a minimum population of 5,000 (Werner & Strosnider, 2020). An evaluation of the Tracking Program subcounty cancer data found that aggregating data across a 5-year span allowed for reliable and stable rates for several cancer types, including lung and colorectal cancers (Ellington et al., 2023).

The Tracking Program cancer incidence rates are based on official federal cancer data from U.S. Cancer Statistics (USCS), which combines cases reported to the National Cancer Institute's Surveillance, Epidemiology and End Results Program and the CDC's National Program of Cancer Registries (Centers for Disease Control and Prevention [CDC], n.d., 2025a, 2025b). We limited our analysis to colorectal and lung cancer incidence because both cancer types were available through the Tracking Network and associated with chemical contaminants detected at Superfund sites in our study area.

We limited the geographic scope of our analysis to the U.S. Census Bureau South Atlantic region because subcounty data for lung cancer and colorectal cancer were available for the majority (75%) of states in this region, including Florida, Georgia, Maryland, North Carolina, Virginia, West Virginia, and Washington, DC. Moreover, because Delaware and South Carolina did not report subcounty cancer data, they were excluded from our analysis. Tracking Network cancer incidence data were suppressed if <16 cases were reported or where there were ≤100 persons in the denominator.

Two distinct measures—the proximity of residents to Superfund sites with chemical risks for colorectal cancer (CRCC) and the proximity of residents to Superfund sites with chemical risks for lung cancer (CRLC)—were calculated for each subcounty area following the methodology described in the U.S. EPA environmental justice mapping tool, EJScreen (U.S. EPA, 2023). These measures reflect the population-average proximity to nearby Superfund sites, with higher scores indicating that residents live closer to sites. We included all current, prior, proposed, and unclassified sites that were within a 5 km buffer of included states and had at least one contaminant that was classified by the U.S. EPA as a contaminant of concern and a risk factor for colorectal cancer or lung cancer. We retained prior sites in our analysis due to the presumption that historic contaminant exposures could impact cancer incidence in later years because of extended latency between onset and diagnosis for many cancer types (Foster et al., 2022; Howard, 2015).

The U.S. EPA classifies chemicals as "contaminants of concern" if they pose an unacceptable risk to human or environmental health based on concentration, toxicology, and potential route of exposure (U.S. EPA, 2025a). Chemicals were classified as cancer risks by cross-referencing U.S. EPA Superfund environmental sampling datasets (U.S. EPA, 2013, 2025b) with 1) the U.S. EPA Integrated Risk Information System, 2) the World Health Organization International Agency for Research on Cancer (IARC), and 3) the National Toxicology Program Chemical Effects in Biological Systems (International Agency for Research on Cancer, 2025; National Toxicology Program, n.d.; U.S. EPA, 2025c). We considered chemicals to be a risk for colorectal or lung cancers if the chemical was designated by one or more of these sources as a known human carcinogen associated with colorectal, lung, or bronchus tumors, excluding any chemicals graded as only a possible or probable human carcinogen. A complete list of all chemicals that were considered a risk factor for lung or colorectal cancer is included in Supplemental Table 1.

Superfund proximity scores were calculated by summing the population-weighted inverse distances between each block group centroid within subcounty areas and all Superfund sites within 5 km (Formula 1) and then dividing by the subcounty population. For block groups with no Superfund sites within 5 km, only the inverse distance between the centroid and the nearest Superfund site was considered. Distances were calculated using the R packages sf (Version 1.0.12) and proxistat (Development Version 1.0.0).

Let X_i = the Superfund proximity score for each subcounty area (i),

$$X_i = \frac{\sum_{j=1}^{n}(p_{j,i}\sum_{k=1}^{m}\frac{1}{d_{j,k}})}{\sum_{j=1}^{n} p_{j,i}} \qquad [1]$$

where $d_{j,k}$ = the distance between each block group (j) centroid and each Superfund site location (k), and $p_{j,i}$ is the population of each block group (j) in subcounty area (i).

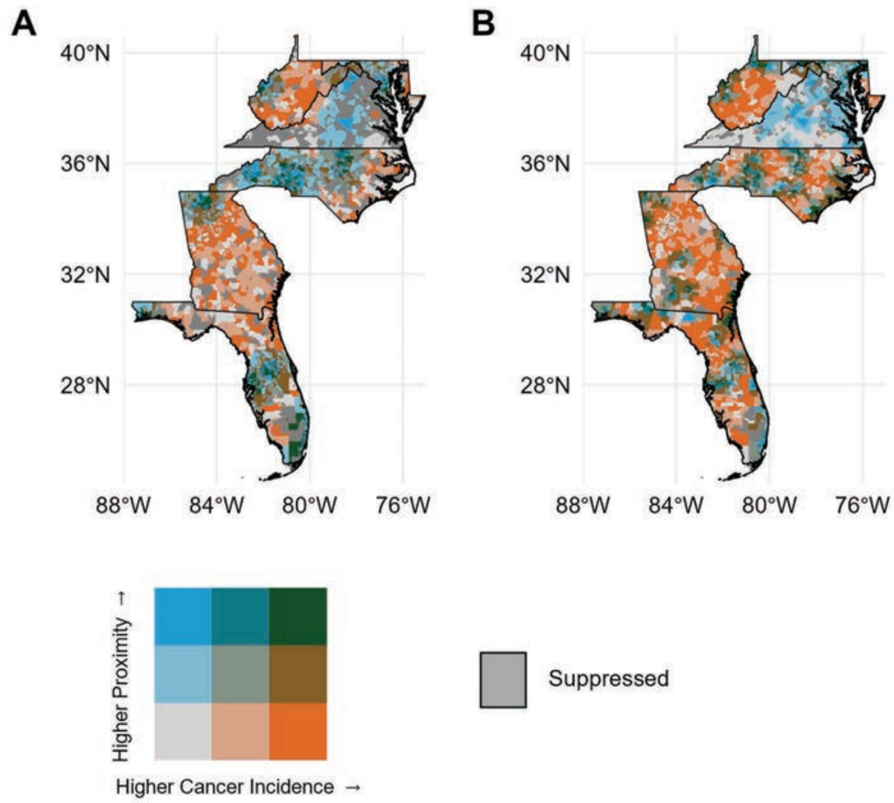
Several additional variables were considered for inclusion in statistical models. We calculated population-weighted average prevalence estimates for behavioral risk factors related to colorectal and lung cancers for each subcounty area using CDC PLACES data (CDC, 2023) and 2014–2018 American Community Survey population estimates. CDC PLACES provides model-based U.S. Census-tract prevalence estimates based on 2018 Behavior Risk Factor Surveillance System data, U.S. Census Bureau 2010 population estimates, and 2014–2018 American Community Survey data (Zhang et al., 2015). Based on prior research, we included CDC PLACES adult prevalence estimates of obesity, current cigarette smoking, no leisure time physical activity (LPA), binge drinking, and colorectal cancer screening among adults ages 50–75 years (Biesalski et al., 1998; Bretthauer et al., 2022; Liu et al., 2019; McNabb et al., 2020; Wolin et al., 2009).

To account for the potential confounding effects of social determinants of health (Tran et al., 2023), we calculated Social Vulnerability Index (SVI) ranks for subcounty areas following the 2018 CDC/ATSDR SVI methodology (Flanagan et al., 2011). The CDC/ATSDR SVI is a composite index that ranks each U.S. Census tract based on four themes of social vulnerability: socioeconomic status, household composition and disability, minority status and language, and housing type and transportation. We calculated both overall and socioeconomic theme ranks for subcounty areas and categorized ranks into quartiles. Due to the high correlation between these values, we included only socioeconomic theme ranks in multivariable analyses.

We estimated the proportion of the subcounty population that resided in an urban area (i.e., percentage of urban-dwelling individuals) by multiplying block group 2014–2018 American Community Survey population estimates by the proportion of the block group land area that overlapped 2010 U.S. Census Bureau urban areas (U.S. Census Bureau, 2011). Block group urban population estimates were then summed by subcounty area, and the sum was divided by the total subcounty area population.

FIGURE **1**

Bivariate Distribution of Colorectal Cancer Incidence Rates and Proximity to Superfund Sites With Chemical Risks for Colorectal Cancer (A) and Lung and Bronchus Cancer Incidence Rates and Proximity to Superfund Sites With Chemical Risks for Lung and Bronchus Cancer (B) by Subcounty Area, South Atlantic Region, 2014–2018



Note. The Centers for Disease Control and Prevention created standardized subcounty areas by spatially aggregating U.S. Census tracts—excluding areas without any resident population—into a contiguous area with a minimum population of 5,000 persons (Werner & Strosnider, 2020). The 13 subcounty areas with suppressed cancer incidence rates are shown in dark gray. Higher proximity scores indicate increased residential proximity to Superfund sites among subcounty residents. Data were not available for South Carolina and Delaware.

Statistical Analysis

We calculated frequencies and proportions of subcounty areas and Superfund sites with CRCC and CRLC by state (Table 1). We visualized geographic patterns in the correlation between cancer incidence rates and Superfund proximity scores using bivariate choropleth maps (Figure 1). We classified subcounty areas with available cancer incidence data by quartile of proximity to Superfund sites with CRCC and CRLC and examined the mean age-adjusted cancer incidence rate and other covariates across quartiles (Table 2).

We used linear multilevel models to characterize the relationship between subcounty age-adjusted incident cancer rates and Superfund proximity scores, with random effects at the state level to account for the clustering of subcounty areas within states. Natural log transformations of dependent variables were performed to improve normality of residuals, as evidenced by visual inspection of model residual Q-Q plots. Regression coefficients and confidence intervals (CIs) represent the relative percent change in the geometric mean of subcounty area cancer incidence rates. A p-value of $\leq .05$ was considered statistically significant.

The final model of log-transformed colorectal cancer incidence rate included Superfund proximity score, average adult obesity and binge drinking prevalences, socioeconomic SVI quartile, and percentage of urban-dwelling individuals. Average no LPA and colorectal cancer screening prevalences were excluded from the adjusted model due to multicollinearity with obesity prevalence and socioeconomic SVI scores, respectively. Multivariable lung cancer incidence models were adjusted for Superfund proximity score, average adult cigarette smoking prevalence, socioeconomic SVI quartile, and percentage of urban-dwelling individuals. Models were constructed using the R package lmer4 (Version 1.1-35.1).

To evaluate potential bias caused by suppression of cancer incidence rates for some subcounty areas, we compared mean values

TABLE **1**

Number and Percentage of Subcounty Areas and Superfund Sites With Chemical Risks for Colorectal and Lung Cancer in the South Atlantic Region by State

	Subcounty Areas * # (%)	Subcounty Area Population # (%)	Superfund Sites With CRCC # (%)	Superfund Sites With CRLC # (%)
Total subcounty areas	6,311 (100)	57,982,008 (100)	21 (100)	181 (100)
Washington, DC	71 (1.1)	684,498 (1.2)	1 (4.8)	1 (0.6)
Florida	2,161 (34.2)	20,598,139 (35.5)	9 (42.9)	67 (37.0)
Georgia	1,126 (17.8)	10,297,484 (17.8)	0 (0)	17 (9.4)
Maryland	684 (10.8)	6,003,435 (10.4)	0 (0)	18 (9.9)
North Carolina	1,122 (17.8)	10,155,624 (17.5)	7 (33.3)	32 (17.7)
Virginia	937 (14.9)	8,413,774 (14.5)	2 (9.5)	31 (17.1)
West Virginia	210 (3.3)	1,829,054 (3.2)	1 (4.8)	7 (3.9)
Neighboring state **	N/A	N/A	1 (4.8)	8 (4.4)

* Excludes 91 subcounty areas in the South Atlantic region that had no resident population.

** Superfund sites in states that neighbored one or more of the included states contributed to the calculation of Superfund proximity if they were within 5 km of the shared border.

Note. Percentages might not add up to 100 due to rounding. The Centers for Disease Control and Prevention created standardized subcounty areas by spatially aggregating U.S. Census tracts—excluding those areas without any resident population—into a contiguous area with a minimum population of 5,000 persons (Werner & Strosnider, 2020). CRCC = chemical risks for colorectal cancer; CRLC = chemical risks for lung cancer; N/A = not applicable.

TABLE **2**

Summary Statistics of Cancer Incidence Rates and Covariates by Proximity Score Quartiles Among Subcounty Areas With Available Cancer Incidence Data, South Atlantic Region, 2014–2018

Summary Statistic	Proximity to Superfund Sites With CRCC Score (*n* = 3,749)				Proximity to Superfund Sites With CRLC Score (*n* = 5,022)			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Quartile 1	Quartile 2	Quartile 3	Quartile 4
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Colorectal cancer incidence rate	49.5 (14.6)	48.3 (15.2)	47.4 (14.0)	49.5 (15.8)	–	–	–	–
Lung and bronchus cancer incidence rate	–	–	–	–	78.4 (24.1)	78.5 (26.1)	73.8 (26.5)	74.4 (27.0)
SVI score	0.57 (0.27)	0.49 (0.28)	0.49 (0.28)	0.57 (0.29)	0.54 (0.28)	0.52 (0.27)	0.44 (0.27)	0.60 (0.28)
SVI socioeconomic status theme score	0.59 (0.27)	0.51 (0.27)	0.49 (0.27)	0.53 (0.29)	0.56 (0.28)	0.54 (0.27)	0.43 (0.27)	0.57 (0.28)
Binge drinking (adults ≥18 years)	15.5 (2.6)	15.6 (2.3)	16.1 (2.5)	16.2 (2.6)	15.6 (2.5)	15.8 (2.4)	16.6 (2.6)	16.2 (2.4)
Current cigarette smoking (adults ≥18 years)	20.9 (4.7)	19.0 (4.8)	18.8 (4.9)	17.1 (4.8)	19.7 (5.0)	20.1 (4.8)	17.6 (4.8)	19.0 (4.6)
Obesity (adults ≥18 years)	34.4 (5.4)	32.9 (5.6)	32.1 (4.8)	31.1 (5.1)	33.1 (6.1)	33.4 (4.7)	31.7 (5.0)	33.1 (5.3)
No leisure-time physical activity (adults ≥18 years)	28.3 (5.5)	26.5 (5.7)	26.0 (5.3)	26.5 (6.8)	27.2 (5.9)	26.8 (5.4)	24.6 (5.4)	27.6 (6.4)
Colorectal cancer screening (adults 50–75 years)	63.7 (5.2)	64.9 (4.6)	65.2 (4.7)	63.8 (6.0)	63.9 (46.8)	58.8 (46.9)	85.7 (32.4)	94.2 (21.9)
Percentage of urban-dwelling population	53.0 (47.9)	73.7 (42.0)	80.7 (37.4)	95.0 (20.3)	62.9 (46.8)	58.8 (46.9)	85.7 (32.4)	94.2 (21.9)

Note. Subcounty areas were grouped into quartiles based on their proximity to Superfund sites with CRCC and CRLC with the first quartile including subcounty areas with residents living furthest from Superfund sites and the fourth quartile including subcounty areas with residents living nearest to Superfund sites. Subcounty areas with suppressed cancer incidence data were excluded from calculations. The Centers for Disease Control and Prevention created standardized subcounty areas by spatially aggregating U.S. Census tracts—excluding areas without any resident population—into a contiguous area with a minimum population of 5,000 persons (Werner & Strosnider, 2020). CRCC = chemical risks for colorectal cancer; CRLC = chemical risks for lung cancer; SVI = Social Vulnerability Index.

TABLE **3**

Linear Multilevel Model for the Association Between Proximity to Superfund Sites and Age-Adjusted Colorectal and Lung Cancer Incidence Rates Among Subcounty Areas, South Atlantic Region, 2014–2018

	Univariate Model		Fully Adjusted Model	
	Estimate (95% CI)	*p*-Value	Estimate (95% CI)	*p*-Value
Colorectal cancer incidence rate model (*n* = 3,749)				
Proximity to Superfund sites with CRCC	35.7 [19.8, 53.8]	<.001	29.2 [15.0, 45.2]	<.001
Obesity prevalence in adults ≥18 years			1.8 [1.5, 2.0]	<.001
Binge drinking prevalence in adults ≥18 years			4.1 [3.7, 4.5]	<.001
SVI socioeconomic status quartile (Reference: Quartile 1)				
Quartile 2			4.5 [1.9, 7.3]	.001
Quartile 3			6.3 [3.3, 9.5]	<.001
Quartile 4			8.9 [5.1, 12.8]	<.001
Percentage of urban-dwelling population			4.0 [1.6, 6.4]	.001
ICC	0.34		0.42	
AIC	1,148.7		570.0	
BIC	1,173.6		632.3	
Lung cancer incidence rate model (*n* = 5,022)				
Proximity to Superfund sites with CRLC	-1.7 [-2.4, 5.9]	.42	-1.9 [-5.3, 1.6]	.28
Smoking prevalence in adults ≥18 years			5.2 [4.9, 5.5]	<.001
SVI socioeconomic status quartile (Reference: Quartile 1)				
Quartile 2			-2.8 [-5.1, -0.5]	.02
Quartile 3			-8.8 [-11.3, -6.3]	<.001
Quartile 4			-21.8 [-24.4, -19.1]	<.001
Percentage of urban-dwelling population			12.1 [9.8, 14.5]	<.001
ICC	0.41		0.47	
AIC	2,619.0		1,022.6	
BIC	2,645.1		1,081.3	

Note. Estimates represent the expected percent difference in the age-adjusted colorectal cancer rate for a 1-unit increase in the covariate. Models include a random intercept for state. The Centers for Disease Control and Prevention created standardized subcounty areas by spatially aggregating U.S. Census tracts—excluding areas without any resident population—into a contiguous area with a minimum population of 5,000 persons (Werner & Strosnider, 2020). AIC = Akaike information criterion; BIC = Bayesian information criterion; CI = confidence interval; CRCC = chemical risks for colorectal cancer; CRLC = chemical risks for lung cancer; ICC = interclass correlation; SVI = Social Vulnerability Index.

of all covariates in subcounty areas with suppressed data to those with available data. Mean differences were tested using a two-sided Student's *t*-test (Supplemental Table 2). All analyses were completed using R (Version 4.3.1).

Results

A total of 6,311 subcounty areas contained or within 5 km of 258 Superfund sites were included in the analysis. CRLC were classified as contaminants of concern at 181 sites (70.2%) and CRCC were classified as contaminants of concern at 21 sites (8.1%) (Table 1). The most common types of chemical contaminants linked to lung cancer were arsenic (126 sites), benzene (107 sites), and cadmium (106 sites). The most common types of chemical contaminants linked to colorectal cancer were bromodichloromethane (19 sites) and bromoform (7 sites) (Supplemental Table 1).

Bivariate maps illustrated areas with both high colorectal cancer incidence rates and high proximity to Superfund sites with CRCC in Northwest Georgia, Western North Carolina, and Central and Southeast Florida (Figure 1). Areas of high rates of lung cancer and proximity to Superfund sites with CRLC did not appear to cluster within regions but were dispersed throughout the region.

Age-adjusted colorectal cancer incidence rates were available for 3,749 (59.4%) subcounty areas, and age-adjusted lung cancer rates were available for 5,022 (79.6%) subcounty areas. Mean age-adjusted colorectal cancer rates were similar across quartiles of proximity to Superfund sites with CRCC, whereas mean age-adjusted lung cancer rates were lowest among subcounty areas in the third and fourth quartiles of proximity to Superfund sites with CRLC (Table 2). Mean overall and socioeconomic SVI scores were lowest in the third quartile of proximity to Superfund sites with CRCC (0.49 and 0.49, respectively) and CRLC (0.44 and 0.43, respectively). The mean prevalence of binge drinking increased slightly with increasing quartile of proximity to Superfund sites with CRCC, whereas the mean prevalence of current cigarette smoking and obesity decreased slightly with increasing quartile of proximity. There were large increases in the percentage of urban-dwelling individuals with increasing quartiles of proximity to Superfund sites with both CRCC and CRLC. The mean percentage of urban-dwelling individuals was 42.0% and 31.3% higher in the fourth quartiles than in the first quartiles of proximity to Superfund sites with CRCC and CRLC, respectively.

Regression models showed a significant univariate association between Superfund site proximity and subcounty colorectal cancer incidence rates (35.7% increase in incidence rate for each 1-unit increase in Superfund proximity score, 95% CI [19.8, 53.8]) but not between Superfund site proximity and subcounty lung cancer incidence rates (1.7% decrease in incidence rate for each 1-unit increase in Superfund proximity score, 95% CI [-2.4, 5.9]) (Table 3).

In the final adjusted model for colorectal cancer incidence, a 1-unit increase in Superfund proximity score was associated with a 29.2% increase (95% CI [15.0, 45.2]) in the incidence rate after adjusting for adult obesity prevalence, adult binge drinking prevalence, socioeconomic SVI score quartile, and percentage of urban-dwelling individuals. In the final adjusted model for lung cancer incidence, Superfund site proximity score had a negative, but not statistically significant, association with subcounty lung cancer incidence rates (1.9% relative decrease, 95% CI [-5.3, 1.6]).

A comparison of subcounty areas with and without cancer incidence revealed that proximity to Superfund sites did not differ by suppression; however, overall SVI scores, socioeconomic SVI scores, average prevalence of current cigarette smoking, obesity, and no LPA were significantly lower in subcounty areas with suppressed data (Supplemental Table 2). The average prevalence of binge drinking and percentage of urban-dwelling individuals were significantly higher among subcounty areas with suppressed colorectal and lung cancer incidence rates.

Discussion

Our analysis presents methods for evaluating associations between residential proximity to Superfund sites and subcounty-level cancer incidence rates in the South Atlantic region using publicly available data provided by the Tracking Network and U.S. EPA. We found that subcounty age-adjusted colorectal cancer rates increased with increasing Superfund proximity score after accounting for potential confounders; however, we did not observe a significant association between Superfund site proximity and age-adjusted lung cancer incidence rates.

Although the methods outlined in our analysis cannot be used to determine a causal link between Superfund proximity and subcounty cancer incidence, our results can inform further explorations of cancer risk factors that affect communities near Superfund sites. Our observation of a statistical association between Superfund proximity and subcounty colorectal cancer incidence rates highlights the importance of considering potential geospatial relationships in cancer investigations. Although the level of inference of our analysis is restricted to population groups, our findings suggest that research involving individuals should also examine the role of environmental exposures in the risk of developing colorectal cancer.

The methods we present in this article could be adapted to explore geospatial relationships across other geographic regions and administrative levels (e.g., STLT health department jurisdictions) or to explore geospatial relationships with other types of environmental hazards. Although we elected to focus on proximity to Superfund sites, researchers could apply these methods to other sources of environmental contamination associated with higher cancer incidence (Fortunato et al., 2011; Howarth & Eiser, 2023). Additionally, these methods could be further enhanced by applying them to smaller geographic units of analysis, where available, to reduce the potential for ecological bias (Tuson et al., 2019). We elected to use subcounty data for this analysis because the data provided the highest available resolution for publicly accessible cancer incidence data across states in the South Atlantic region. Although subcounty areal units allow for more precise classification of exposures and outcomes than counties or states, they still could be subject to ecological bias resulting from arbitrary aggregation along U.S. Census-tract boundaries. Therefore, STLT health departments that have access to cancer incidence data for smaller geographic units could reduce the potential for bias by applying the described methods to the smallest geographic unit available.

Our methods have several limitations to consider. First, we used a singular measure of exposure (i.e., proximity to Superfund sites) to ascribe risk to subcounty populations, although the exposure risk could vary depending on the types of chemical contaminants present, the type of media (e.g., soil, water, air) contaminated, and the extent to which contaminants affect exposure pathways. Public data on Superfund site contaminants do not contain enough information to accurately characterize exposure pathways for the population residing near the site, and the detection of chemical contaminants alone does not confirm that any individuals were exposed. The duration of potential exposures within communities might also vary. Although U.S. EPA provides historical site assessment and response information, it is not possible to create an accurate timeline for each site due to the lack of uniform data on when contamination first occurred. Contamination might be present for years before discovery, leading to uncertainty about the duration of exposure. Further, characterization of exposure duration is constrained by lack of information about residential history.

For example, the population living near a Superfund site might include individuals who have recently arrived or exclude individuals who have recently moved away. Due to the long latency period for some types of cancer, consideration of both site and residential histories is necessary to accurately ascertain exposure status. Investigations into the relationship between specific Superfund sites and cancer incidence can be strengthened by considering the duration and concentration of contaminants in the exposure pathways most likely to affect cancer risk.

Second, additional unmeasured, individual-level factors such as usual water sources and contemporaneous exposures could confound or interact with the relationship between proximity to Superfund sites and cancer incidence. For example, long-term air pollution levels; occupational exposures to asbestos, silica, metals, and polycyclic aromatic hydrocarbons; dietary habits; genetic risk factors; and ionizing radiation exposure are important contributors to lung cancer risk (Malhotra et al., 2016); however, we were not able to control for these factors due to the unavailability of reliable data at the subcounty level. Although our analysis controlled for the prevalence of important colorectal and lung cancer risk factors (e.g., obesity, smoking prevalence) in the subcounty population, residual confounding factors might remain because we were unable to control for joint covariate distributions (Greenland & Robins, 1994).

Third, suppression of cancer incidence data for subcounty areas with a low number of cases could have influenced our findings, particularly for our analysis of colorectal cancer incidence rates, which were suppressed in approximately 40% of subcounty areas. We attempted to minimize information bias by controlling for factors that varied by data availability in statistical models; however, the extent to which this action controls for selection bias depends on identification of relevant variables, accuracy of measurement, complexity of selection mechanism, and unobserved confounders (Hernán & Robins, 2020).

Despite these limitations, our study provides a replicable method for analyzing the distribution of environmental hazards and health outcomes across populations using widely available public data. These methods can be used to support proactive monitoring of cancer patterns or as an initial step in the cancer investigation process.

Conclusion

Among subcounty areas in the South Atlantic region, we observed higher rates of colorectal cancer incidence with increasing residential proximity to Superfund sites that had prior contamination with chemicals known to increase the risk for colorectal cancer, but we did not find an association between lung cancer incidence and residential proximity to Superfund sites. The methods described in our study can be adapted to reflect the unique environmental and health concerns of local communities. These analyses demonstrate how public data on environmental hazards and population demographics can be leveraged to strengthen core cancer monitoring and evaluation activities.

Disclaimer: The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of CDC or ATSDR. The authors received no external financial support for the research, authorship, or publication of this article. The authors declare they have no competing interests related to this work to disclose. This study used no individual-level data and therefore was exempt from IRB review, was reviewed by CDC, and was conducted by adhering to applicable federal law and CDC policy.

Acknowledgments: The authors thank Drs. June K. Dunnick, Darlene Dixon, and Shanshan Zhou of the National Institutes for Environmental Health Sciences for reviewing and providing commentary on the manuscript.

Corresponding Author: Sarah Rockhill, MPH, Geospatial Epidemiologist, Geospatial Research, Analysis, and Services Program, Agency for Toxic Substances and Disease Registry, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA 30341. Email: ppk8@cdc.gov

References

Balluz, L.S. (2014). CDC's Environmental Public Health Tracking Network: An innovative dynamic surveillance system for you. *Journal of Environmental Health*, *76*(7), 48–51.

Biesalski, H.K., De Mesquita, B.B., Chesson, A., Chytil, F., Grimble, R., Hermus, R.J.J., Köhrle, J., Lotan, R., Norpoth, K., Pastorino, U., & Thurnham, D. (1998). European consensus statement on lung cancer: Risk factors and prevention. Lung Cancer Panel. *CA: A Cancer Journal for Clinicians*, *48*(3), 167–176. https://doi.org/10.3322/canjclin.48.3.167

Bretthauer, M., Løberg, M., Wieszczy, P., Kalager, M., Emilsson, L., Garborg, K., Rupinski, M., Dekker, E., Spaander, M., Bugajski, M., Holme, Ø., Zauber, A.G., Pilonis, N.D., Mroz, A., Kuipers, E.J., Shi, J., Hernán, M.A., Adami, H.-O., Regula, J., . . . Kaminski, M.F. (2022). Effect of colonoscopy screening on risks of colorectal cancer and related death. *The New England Journal of Medicine*, *387*(17), 1547–1556. https://doi.org/10.1056/NEJMoa2208375

Centers for Disease Control and Prevention. (n.d.). *2014–2018 age-adjusted rate of lung and colorectal cancer per 100,000 population (5,000 min. population area, 5-year period)*. https://ephtracking.cdc.gov/DataExplorer/

Centers for Disease Control and Prevention. (2023). *PLACES: Census tract data (GIS friendly format), 2020 release*. https://data.cdc.gov/500-Cities-Places/PLACES-Census-Tract-Data-GIS-Friendly-Format-2020-/ib3w-k9rq/about_data

Centers for Disease Control and Prevention. (2025a). *Indicators and data. Indicator: Incidence of bladder, brain, kidney, leukemia, liver, lung, melanoma, non-Hodgkin lymphoma and thyroid cancer*. https://ephtracking.cdc.gov/indicatorPages

Centers for Disease Control and Prevention. (2025b). *Indicators and data. Indicator: Incidence of colorectal cancer*. https://ephtracking.cdc.gov/indicatorPages

Currie, J., Greenstone, M., & Moretti, E. (2011). Superfund cleanups and infant health. *American Economic Review*, *101*(3), 435–441. https://doi.org/10.1257/aer.101.3.435

Ellington, T.D., Werner, A.K., Henley, S.J., Paddock, L.E., & Agovino, P.K. (2023). Feasibility of visualizing cancer incidence data at sub-county level: Findings from 21 National Program of Cancer Registries. *Spatial and Spatio-temporal Epidemiology*, *45*, Article 100564. https://doi.org/10.1016/j.sste.2023.100564

Flanagan, B.E., Gregory, E.W., Hallisey, E.J., Heitgerd, J.L., & Lewis, B. (2011). A social vulnerability index for disaster management. *Journal of Homeland Security and Emergency Management*, *8*(1), Article 3. https://doi.org/10.2202/1547-7355.1792

Fortunato, L., Abellan, J.J., Beale, L., LeFevre, S., & Richardson, S. (2011). Spatio-temporal patterns of bladder cancer incidence in Utah (1973–2004) and their association with the presence of toxic release inventory sites. *International Journal of Health Geographics*, *10*, Article 16. https://doi.org/10.1186/1476-072X-10-16

References

Foster, S.L., Lavery, A.M., Condon, S.K., Etheredge, A.A., Kennedy, B.S., Svendsen, E.R., & Breysse, P. (2022). *Guidelines for examining unusual patterns of cancer and environmental concerns*. Centers for Disease Control and Prevention. https://www.cdc.gov/cancer-environment/php/guidelines/index.html

Greenland, S., & Robins, J. (1994). Invited commentary: Ecologic studies—Biases, misconceptions, and counterexamples. *American Journal of Epidemiology*, *139*(8), 747–760. https://doi.org/10.1093/oxfordjournals.aje.a117069

Hernán, M.A., & Robins, J.M. (2020). *Causal inference: What if*. Chapman & Hall/CRC Press.

Howard, J. (2015). *Minimum latency & types or categories of cancer*. World Trade Center Health Program. https://www.cdc.gov/wtc/pdfs/policies/WTCHP-Minimum-Cancer-Latency-PP-01062015-508.pdf

Howarth, M.V., & Eiser, A.R. (2023). Environmentally mediated health disparities. *The American Journal of Medicine*, *136*(6), 518–522. https://doi.org/10.1016/j.amjmed.2023.02.008

International Agency for Research on Cancer. (2025). *List of classifications by cancer sites with sufficient or limited evidence in humans, IARC Monographs Volumes 1–139*. World Health Organization. https://monographs.iarc.who.int/wp-content/uploads/2019/07/Classifications_by_cancer_site.pdf

Iyer, R., Aggarwal, J., & Iken, B. (2016). A review of the Texas, USA San Jacinto Superfund site and the deposition of polychlorinated dibenzo-p-dioxins and dibenzofurans in the San Jacinto River and Houston Ship Channel. *Environmental Science and Pollution Research*, *23*, 23321–23338. https://doi.org/10.1007/s11356-016-7501-8

Kiaghadi, A., Rifai, H.S., & Dawson, C.N. (2021). The presence of Superfund sites as a determinant of life expectancy in the United States. *Nature Communications*, *12*(1), Article 1947. https://doi.org/10.1038/s41467-021-22249-2

Liu, P.-H., Wu, K., Ng, K., Zauber, A.G., Nguyen, L.H., Song, M., He, X., Fuchs, C.S., Ogino, S., Willett, W.C., Chan, A.T., Giovannucci, E.L., & Cao, Y. (2019). Association of obesity with risk of early-onset colorectal cancer among women. *JAMA Oncology*, *5*(1), 37–44. https://doi.org/10.1001/jamaoncol.2018.4280

Malhotra, J., Malvezzi, M., Negri, E., La Vecchia, C., & Boffetta, P. (2016). Risk factors for lung cancer worldwide. *European Respiratory Journal*, *48*(3), 889–902. https://doi.org/10.1183/13993003.00359-2016

McNabb, S., Harrison, T.A., Albanes, D., Berndt, S.I., Brenner, H., Caan, B.J., Campbell, P.T., Cao, Y., Chang-Claude, J., Chan, A., Chen, Z., English, D.R., Giles, G.G., Giovannucci, E.L., Goodman, P.J., Hayes, R.B., Hoffmeister, M., Jacobs, E.J., Joshi, A.D., . . . Peters, U. (2020). Meta-analysis of 16 studies of the association of alcohol with colorectal cancer. *International Journal of Cancer*, *146*(3), 861–873. https://doi.org/10.1002/ijc.32377

National Toxicology Program. (n.d.). *Chemical effects in biological systems*. National Institute of Environmental Health Sciences. https://cebs-ext.niehs.nih.gov/cebs/

Office of Land and Emergency Management. (2020). *Population surrounding 1,857 Superfund remedial sites*. U.S. Environmental Protection Agency. https://www.epa.gov/sites/default/files/2015-09/documents/webpopulationrsuperfundsites9.28.15.pdf

Press, D.J., McKinley, M., Deapen, D., Clarke, C.A., & Gomez, S.L. (2016). Residential cancer cluster investigation nearby a Superfund Study Area with trichloroethylene contamination. *Cancer Causes & Control*, *27*(5), 607–613. https://doi.org/10.1007/s10552-016-0734-5

Pugh, K.H., & Zarus, G.M. (2012). The burden of environmental disease in the United States. *Journal of Environmental Health*, *74*(9), 30–34.

Tran, T., Rousseau, M.A., Farris, D.P., Bauer, C., Nelson, K.C., & Doan, H.Q. (2023). The social vulnerability index as a risk stratification tool for health disparity research in cancer patients: A scoping review. *Cancer Causes & Control*, *34*(5), 407–420. https://doi.org/10.1007/s10552-023-01683-1

Tuson, M., Yap, M., Kok, M.R., Murray, K., Turlach, B., & Whyatt, D. (2019). Incorporating geography into a new generalized theoretical and statistical framework addressing the modifiable areal unit problem. *International Journal of Health Geographics*, *18*(1), Article 6. https://doi.org/10.1186/s12942-019-0170-3

U.S. Census Bureau. (2011). Urban area criteria for the 2010 Census. *Federal Register*, *76*(164), 53030–53043. https://www.federalregister.gov/documents/2011/08/24/2011-21647/urban-area-criteria-for-the-2010-census

U.S. Environmental Protection Agency. (2011). *This is Superfund: A community guide to EPA's Superfund program*. https://semspub.epa.gov/work/HQ/175197.pdf

U.S. Environmental Protection Agency. (2013). *Contaminants at CERCLIS sites (List-10), Version 1.02*. https://semspub.epa.gov/work/HQ/100000010.pdf

U.S. Environmental Protection Agency. (2023). *Environmental justice mapping and screening tool: EJScreen technical documentation for Version 2.2*. https://downloads.regulations.gov/EPA-HQ-OW-2021-0736-0434/content.pdf

U.S. Environmental Protection Agency. (2025a). *About risk assessment*. https://www.epa.gov/risk/about-risk-assessment

U.S. Environmental Protection Agency. (2025b). *Contaminant of concern data for decision documents by media, FYs 1981–2024*. https://www.epa.gov/superfund/superfund-data-and-reports

U.S. Environmental Protection Agency. (2025c). *Integrated Risk Information System*. https://www.epa.gov/iris

Werner, A.K., & Strosnider, H.M. (2020). Developing a surveillance system of sub-county data: Finding suitable population thresholds for geographic aggregations. *Spatial and Spatio-tem-*

continued on page 26

References *continued from page 25*

poral Epidemiology, *33*, Article 100339. https://doi.org/10.1016/j.sste.2020.100339

Wolin, K.Y., Yan, Y., Colditz, G.A., & Lee, I.-M. (2009). Physical activity and colon cancer prevention: A meta-analysis. *British Journal of Cancer*, *100*(4), 611–616. https://doi.org/10.1038/sj.bjc.6604917

Zhang, X., Holt, J.B., Yun, S., Lu, H., Greenlund, K.J., & Croft, J.B. (2015). Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the behavioral risk factor surveillance system. *American Journal of Epidemiology*, *182*(2), 127–137. https://doi.org/10.1093/aje/kwv002

Zhang, Y., McDermott, S., Davis, B., & Hussey, J. (2020). High incidence of brain and other nervous system cancer identified in two mining counties, 2001–2015. *Spatial and Spatio-temporal Epidemiology*, *32*, Article 100320. https://doi.org/10.1016/j.sste.2019.100320

Did You Know?

Our Environmental Public Health Emergency Preparedness and Response (EPHEPR) capability framework is available to help state, tribal, local, and territorial jurisdictions better assess and strengthen their readiness to address core public health issues in disasters. As a guidance resource, the framework supports public health agencies in formally integrating environmental public health into their emergency preparedness, response, and recovery planning and coordination activities within the health agency. Learn more at www.neha.org/epr-framework.




Now Available!

Updated Registered Environmental Health Specialist/Registered Sanitarian (REHS/RS) Study Guide, 5th Edition

- Fresh visual layout to enhance reading and studying experience
- 15 chapters covering critical exam content
- Insights from 29 experts

Helps you identify where to focus your studying so you can pass the exam!

neha.org/rehs-study-materials

neha National Environmental Health Association

Copyright 2025, National Environmental Health Association (www.neha.org)