air

MDPI

Article

Total Toxic Releases from Electric Utilities and Mining Facilities and Their Relationships with Human Health in the United States

Azita Amiri [1,*], Xiaoxia Dong [2], Armita Amiri [3], Shuang Zhao [4] and Mary Fox [5]

[1] College of Nursing, University of Alabama in Huntsville, Huntsville, AL 35899, USA
[2] Weitzman School of Design, University of Pennsylvania, Philadelphia, PA 19104, USA; xiaoxiad@design.upenn.edu
[3] College of Arts and Sciences, Florida State University, Tallahassee, FL 32306, USA; armitaamiri2020@gmail.com
[4] Political Science Department, University of Alabama in Huntsville, Huntsville, AL 35899, USA; sz0009@uah.edu
[5] Risk Sciences and Public Policy Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA; mfox9@jhu.edu
* Correspondence: azita.amiri@uah.edu

Abstract

This manuscript examines the total toxic releases from electric utilities and mining facilities in the United States in 2020, focusing on their relationships with human health outcomes. The research highlights the adverse effects of air and water pollution, linking exposure to toxic emissions to several health issues, such as low birth weight, respiratory and cardiovascular diseases, and cancers. It underscores the disproportionate impact of these pollutants on low-income and minority populations. This research project utilizes two sets of data: (1) environmental data, the EPA's Toxic Release Inventory (TRI) data that records total emissions of all-electric and mining facilities, and (2) health data, the PLACES health data. The results of this study show that census tracts exposed to higher toxic releases are expected to have worse health outcomes. The coefficients for total toxic release indicate that higher toxic release corresponds to a higher rate of cancer, chronic obstructive pulmonary disease (COPD), diabetes, kidney diseases, arthritis, cardiac heart disorders (CHD), and stroke. Except for diabetes and kidney diseases, the associations are statistically significant. The analysis indicates the need for comprehensive public health strategies to mitigate the risks posed by toxic releases, particularly for vulnerable communities.

Keywords: total toxic release; TRI; electric facilities; mining facilities; health; chronic disorders; cancer; COPD; PLACES; stroke

Academic Editors: Nedim Durmus and Ling Tim Wong

Received: 17 February 2026
Revised: 11 May 2026
Accepted: 12 May 2026
Published: 14 May 2026

1. Introduction

Research examining the health consequences of environmental pollution has consistently demonstrated that exposure to industrial air and water contaminants is associated with adverse health outcomes. A large body of epidemiologic evidence links ambient pollution and toxic chemical exposure to respiratory diseases, cardiovascular disease, metabolic disorders, kidney diseases, cancer, and adverse birth outcomes [1–7]. These impacts are not evenly distributed, as numerous studies have documented that low income and racially marginalized communities are disproportionately exposed to industrial emissions and pollution burdens [8–11].

Industries and businesses in the United States (U.S.) use many chemicals to make the products we depend on. Also, electric utilities generate, transmit, and distribute

electric power. These facilities are subject to reporting to the United States Environmental Protection Agency's (EPA) Toxics Release Inventory (TRI) program. TRI collects data on some chemicals released into the environment. While many chemical releases are reported to the TRI, the TRI program does not cover all chemicals released by industries, for instance, most greenhouse gas emissions are not reported to the TRI [12,13]. The TRI collects chemical releases in units of pounds, except for dioxin and dioxin-like compounds, which are reported in grams. The TRI compiles data on how facilities release or dispose of chemical waste into the environment, including by directly releasing it into the air or water, on land, or by transferring it to an off-site location. The data reported to the EPA are publicly available. For the calendar year 2020, more than 21,000 facilities reported to EPA's TRI Program [12].

Electric utilities, coal mining, and metal mining are included in TRI. For 2020, 3463 electric utilities, 716 metal mining, and 94 coal mining generating facilities reported their environmental waste to the TRI [12]. These utilities use a variety of fuels, including fossil natural gas and nuclear power, to generate electricity. Electric utilities and coal and metal mining facilities can add heavy metals, such as arsenic, chromium, cobalt, lead, and nickel, and other chemicals, such as benzene, aldehydes, and polycyclic aromatic compounds to the environment [14]. Some of these chemicals, such as lead, arsenic, formaldehyde, and benzene, are known as human carcinogens, and some affect the respiratory system or other organs [15–18].

Human health risks from chemical exposure are determined by many factors, including chemical type and toxicity, quantity, environmental media, genetics, and other sources of exposure (e.g., food, water, consumer products) [13]. Therefore, the amount of chemical releases by itself is not necessarily an indicator of human risk. In addition, TRI data generally cannot imply the extent of exposure to chemicals. However, further analysis of TRI data can be used to assess potential exposure and human health risks. The EPA has developed relevant tools, such as the Risk-Screening Environmental Indicators (RSEI) model and NATA Modeled Ambient Concentrations, Exposures, and Risks, which provide researchers with information on the potential hazards of released chemicals and the total TRI releases [14,19,20].

Although release and disposal practices are subject to various regulatory requirements and restrictions and are designed to minimize potential exposures or harm to human health and the environment, several studies have shown the negative impacts of these toxic releases on human health [1,21,22]. During the COVID-19 pandemic, counties with more Toxic Release Inventory (TRI) sites had a 53% increase in COVID-19 cases and a 10.6% increase in COVID-19 deaths [23].

Prior research on industrial emissions and health outcomes can be broadly categorized into three overlapping pillars. First, toxicological and epidemiologic studies have linked specific pollutants, such as heavy metals, volatile organic compounds (VOCs), and combustion byproducts, to biological mechanisms including oxidative stress, systemic inflammation, endocrine disruption, and DNA damage that underline chronic disease development [15–18,24–26]. Second, environmental justice scholarship has shown persistent racial and socioeconomic disparities in proximity to polluting facilities and cumulative exposure burdens [8–11]. Third, methodological studies have leveraged regulatory datasets, particularly the TRI, to characterize spatial patterns of industrial emissions and potential population-level health risks [10,21,22].

Although numerous studies have examined associations between industrial pollution and adverse health outcomes [1,7], important gaps remain in the scale, resolution, and sector-specific focus of existing research. To our knowledge, few national-level studies have simultaneously focused on electric utilities and mining facilities at the census-tract

level across the United States, despite these sectors accounting for a substantial share of TRI-reported releases and emitting pollutants with well-established biological relevance to chronic diseases. Two sectors that together account for more than half of all TRI-reported releases and contribute disproportionately to heavy metal and combustion-related emissions. Additionally, most studies have relied on county-level or regional data, limiting their ability to capture smaller-scale community exposure patterns. Taken together, prior studies demonstrate strong links between industrial pollution, chronic disease outcomes, and environmental disparities. However, existing research has largely relied on county-level or regional analyses, focused on single industries, or lacked integration with nationally comparable tract level health data. This study addresses these gaps by linking census tract level TRI emissions from electric utility and mining facilities with PLACES health indicators across the entire United States, enabling a high-resolution assessment of population health disparities associated with major industrial emission sources.

By integrating tract-level TRI data with the CDC PLACES dataset, this study provides a high-resolution national assessment of chronic disease patterns associated with industrial emissions. The application of a state-fixed-effects modeling framework further strengthens inference by isolating within-state variation and accounting for unobserved regulatory, policy, and health system differences across states. The inclusion of interaction terms capturing racial composition across urban and rural areas further contributes to a novel environmental justice perspective that has been largely absent from prior regional or state-level analyses.

This study uses an ecological design. Ecological study designs are widely used in environmental epidemiology to examine associations between area-level exposures and population health outcomes. While such designs cannot establish individual-level causality, they are particularly valuable for identifying large-scale spatial disparities, informing regulatory priorities, and generating hypotheses for more detailed exposure studies. Prior studies using TRI data have documented associations with cancer incidence, adverse birth outcomes, and respiratory morbidity, often at regional or county scales [1,21,22]. However, these studies frequently aggregate data over large geographic areas or focus on single industries or regions. Ecological studies cannot infer individual-level risk, but are valuable for identifying large-scale spatial patterns and environmental justice concerns.

By applying a state fixed effects modeling framework and explicitly examining urban–rural racial composition interactions, this study contributes a national, high resolution assessment of environmental health disparities associated with major industrial sectors. The objective is not to infer individual causation but to identify population-level patterns consistent with environmental justice concerns and biologically plausible exposure–disease relationships.

The primary aim of this study was to examine the association between census tract-level toxic releases from electric utility and mining facilities and the prevalence of selected chronic diseases across the United States. We hypothesize that census tracts with higher levels of toxic releases will exhibit a higher prevalence of rates of chronic diseases, including cancer, COPD, coronary heart disease, arthritis, and stroke. We further hypothesize that these associations were stronger in rural communities and in census tracts with higher proportions of non-White residents, reflecting underlying environmental justice disparities.

2. Materials and Methods

This study employed a cross-sectional ecological design using census tract-level data for the year 2020. Data on industrial emissions were obtained from the EPA's TRI, while population health outcomes were derived from the CDC PLACES dataset. These datasets were spatially linked at the census tract level, and fixed effects linear regression models

with state indicators were used to examine associations between toxic releases and chronic disease prevalence while adjusting for sociodemographic and geographic characteristics.

This research project utilized two sets of data: (1) environmental data, the EPA's Toxic Release Inventory (TRI) data that records the total emissions of all electric and mining facilities [27], and (2) health data, the PLACES health data, a database that is a production of collaboration between CDC, the Robert Wood Johnson Foundation, and the CDC Foundation [28]. The PLACES database provides information on cases of different diseases at different geographic levels. We selected eight diseases linked to environmental pollution. We selected both datasets at the tract level to merge them and systematically answer our research question: whether exposure to higher levels of pollution leads to more sickness. The PLACES data have evolved in recent years, with 2020 offering the most comprehensive samples [29]. The PLACES dataset has been available since 2016. The 2016–2019 PLACES data include 27,210 data points, but in 2020, the number increased to 72,337.

Because the CDC PLACES dataset provides complete national census-tract coverage only beginning in 2020, this study used a cross-sectional ecological design focused on that year to maximize spatial completeness and alignment between health and data. While reliance on a single year limits the ability to examine cumulative exposures or temporal trends, the 2020 dataset enables a uniquely comprehensive national tract-level analysis. We acknowledge that health estimates for 2020 may also reflect indirect effects of the COVID-19 pandemic, which could introduce unmeasured confounding. This limitation is addressed in the discussion.

Furthermore, PLACES health estimates include only adults aged 18 and older, which limits interpretation for conditions with substantial childhood prevalence, such as asthma. Individual-level behavioral risk factors (e.g., smoking, diet, occupation, and healthcare access) are unavailable in this ecological design, and residual confounding is therefore expected.

TRI data have well-recognized limitations. The TRI excludes several important pollutants, including most greenhouse gases and fine particulate matter (PM2.5), and relies on facility self-reporting or modeled estimates rather than direct measurements. This approach captures the overall industrial burden rather than individual exposure. These limitations may lead to the underestimation or misclassification of true exposure burdens. Furthermore, using the total mass of release as the exposure metric does not account for differences in chemical toxicity, persistence, or dispersion. While this approach captures the overall industrial burden, it may obscure the health relevance of specific pollutants. Figure 1 shows the distribution of electric utilities, coal mining, and metal mining facilities in the US.

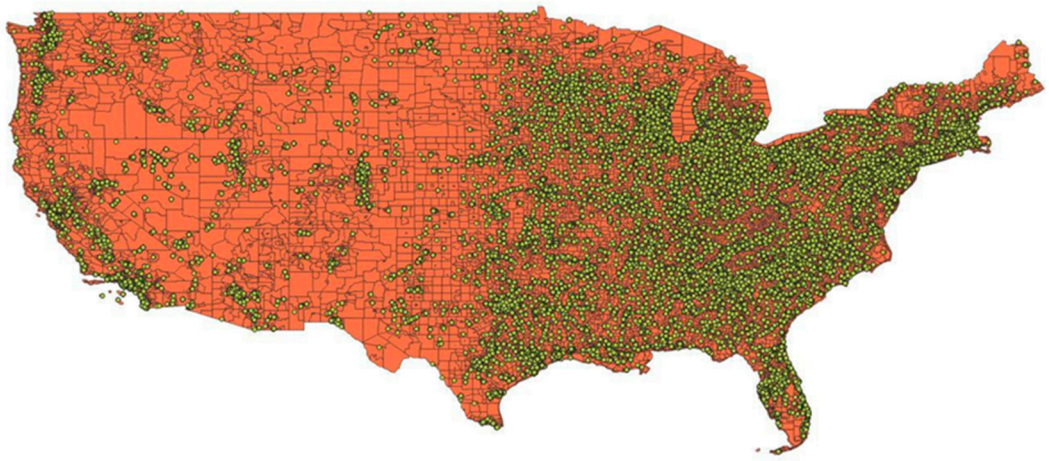

Figure 1. Distribution of electric utilities, coal mining, and metal mining facilities in the U.S.

2.1. Dependent and Independent Variables

The eight disease types served as our dependent variable. The PLACES database provides prevalent data for each disease. We estimated eight models, with each disease prevalence rate serving as the dependent variable. The eight diseases included in our models were: cancer, asthma, COPD, diabetes, kidney disease, arthritis, CHD, and stroke. No transformation of variables was conducted. We directly used the prevalence, the number of cases in every one thousand persons, sometimes referred to as the prevalence rate, which made the data comparable across jurisdictions without controlling for population size (e.g., larger population is inherently associated with a higher case rate).

The main independent variable was the total TRI emission from U.S. electricity and mining facilities. The unit of measurement was in tons. We first transformed the variable to million tons and then took its log to accommodate a nonlinear distribution of the independent variable.

Exposure was operationalized as the log-transformed total mass (tons) of TRI reported releases from electric utility and mining facilities located within each census tract. This measure represents an indicator of industrial emission burden rather than a direct estimate of individual-level chemical exposure. This approach does not differentiate pollutants by toxicity, persistence, or route of exposure and does not incorporate atmospheric or hydrologic dispersion modeling. As a result, tracts adjacent to emitting facilities may experience exposure misclassification, and releases may affect neighboring tracts beyond administrative boundaries. Additionally, the total mass of releases may understate the health relevance of substances with smaller volumes but greater toxicity.

Despite these limitations, mass-based TRI measures are commonly used in national ecological analyses to identify population-level patterns and environmental justice concerns, particularly when toxicity-weighted metrics or measured ambient concentrations are not consistently available at fine spatial resolution. Accordingly, the exposure metric in this study is intended to capture relative industrial emission burden across communities rather than precise individual exposure.

2.2. Controls

Many factors may affect a person's likelihood of developing any disease, including socioeconomic background. Research has also suggested that different races may have different susceptibility to certain diseases. Similarly, these differences may apply to men and women. The literature suggests that a more educated population tends to be more proactive in their health, taking measures such as early screening to prevent certain diseases. Hence, we included race, sex, household income, and education as our control variables. We used U.S. Census data from 2018 in our model as the closest match to our dependent and independent variables by year. Additionally, because this is an ecological study, individual level risk factors such as smoking, diet, occupational exposures, and healthcare access could not be controlled, and residual confounding is likely.

2.3. Data Analysis Strategy

We estimated linear fixed-effects regression models, which allowed us to control for unobserved characteristics that are constant within each state but vary across states. A fixed effects model removes all time-invariant differences between states, such as long standing differences in health systems, regulatory environments, or environmental enforcement, so that the estimated coefficients reflect associations within states rather than across them. In other words, these unobserved state-level effects, such as public health policy and resources, are controlled by fixed effects. In practice, this is implemented by including a set of state indicator (dummy) variables in each model. These indicators absorb any unmeasured, state-

specific factors that could otherwise bias the relationship between toxic releases and health outcomes. Because our analysis used 2020 cross-sectional data, the fixed effects operate across geographic units rather than across time. This approach strengthens internal validity by ensuring that comparisons are made among census tracts within the same state, holding constant all state-level characteristics. A fixed effects model is estimated as follows:

$$Y_i = \chi X_i + \delta \nu_i + u_i$$

where ν_i represents the unit effects.

As mentioned earlier, because this study is an ecological design, individual-level behavioral and clinical risk factors such as smoking, diet, occupational exposures, physical activity, and healthcare access were not available and could not be directly controlled. Sociodemographic variables included in the models (poverty, education, race, sex, age, and urban/rural topology) serve as partial proxies for underlying health risk environments but do not eliminate the possibility of residual confounding.

Consequently, estimated associations between toxic releases and health outcomes should be interpreted as population-level correlations, and not causal effects at the individual level. This limitation is inherent to ecological studies and underscores the role of this analysis as a hypothesis-generating and policy-relevant assessment of spatial health disparities.

3. Results

Table 1 presents the output from multilevel models examining health outcomes, toxic releases, and a series of socio-economic and geographic variables. The coefficients reported in Table 1 are linear regression coefficients representing the expected change in the crude prevalence of each disease (cases per 1000 adults) associated with a one unit change in each predictor, holding all other variables constant. For the main independent variable (log total TRI release), the coefficients represent the change in disease prevalence associated with a proportional change in toxic releases. Positive coefficients indicate that higher values of the predictor are associated with higher disease prevalence, whereas negative coefficients indicate a negative association with lower disease prevalence. These coefficients should not be interpreted as relative risks or odds ratios; they reflect absolute changes in prevalence rates within the linear fixed effects framework.

The coefficients for log-transformed total toxics release indicate the change in health outcomes for each percentage change in total toxics release. For example, for each 10% increase in total toxic releases, the crude rate of cancer in a census tract where the pollution source facility is located is expected to increase by approximately 0.0004 percentage points, holding all other predictors constant. This was calculated by multiplying the coefficient for log total TRI in Model 1, 0.004, by the natural log of 1.1 (i.e., 1 + 10%). For the socio-economic variables, the coefficients have direct interpretations as changes in health outcomes associated with a one-unit change in the socio-economic variables. For example, a 1 percentage point increase in the share of residents living in poverty in a census tract corresponds to an average increase of roughly 0.004 percentage points in the crude cancer rate, all else being equal. In the analysis, the West Census region was treated as the reference category. That is, the Midwest, Northeast, and South Census regions were compared to the West region. Additionally, census tracts located in urban areas, as defined by the Census Bureau, were treated as the reference category and compared to census tracts in rural areas. The interaction between census tract location in urban/rural areas and the percentage of non-White residents revealed distinct associations between health outcomes and non-White residents across urban and rural census tracts.

Table 1. Fixed-effects regression coefficients estimating the associations between total toxic releases and crude disease prevalence rates across U.S. Census tracts.

Diseases Models	Cancer (1)	Asthma (2)	COPD (3)	Diabetes (4)	Kidney Disorders (5)	Arthritis (6)	CHD (7)	Stroke (8)
				Coefficient (S.E.)				
Intercept	−3.279 *** (0.084)	7.367 *** (0.191)	−0.017 (0.250)	−2.434 *** (0.219)	−0.911 *** (0.035)	−4.408 *** (0.633)	−2.646 *** (0.139)	−2.078 *** (0.061)
Total TRI (log)	0.004 *** (0.001)	−0.005 *** (0.001)	0.007 *** (0.001)	0.001 (0.002)	0.001 (0.001)	0.006 (0.004)	0.006 *** (0.001)	0.003 *** (0.001)
Non-White (%)	−0.016 *** (0.0002)	0.013 *** (0.0002)	−0.011 *** (0.0002)	0.052 *** (0.0004)	0.007 *** (0.0001)	−0.025 *** (0.001)	−0.008 *** (0.0002)	0.014 *** (0.0001)
In poverty (%)	0.004 *** (0.000)	0.060 *** (0.0003)	0.104 *** (0.0005)	0.130 *** (0.001)	0.037 *** (0.0002)	0.132 *** (0.001)	0.085 *** (0.0004)	0.051 *** (0.0003)
% female (%)	0.066 *** (0.001)	0.039 *** (0.001)	0.037 *** (0.001)	0.062 *** (0.001)	0.023 *** (0.0004)	0.199 *** (0.003)	0.043 *** (0.001)	0.029 *** (0.001)
Bachelor's degree or higher (%)	−0.006 *** (0.0002)	−0.027 *** (0.0002)	−0.067 *** (0.0003)	−0.084 *** (0.0004)	−0.017 *** (0.0001)	−0.126 *** (0.001)	−0.047 *** (0.0003)	−0.026 *** (0.000)
Median age	0.179 *** (0.001)	0.0001 (0.0004)	0.133 *** (0.001)	0.224 *** (0.001)	0.063 *** (0.0002)	0.536 *** (0.002)	0.180 *** (0.001)	0.088 *** (0.0003)
Typology (reference: urban)	−0.125 *** (0.011)	0.225 *** (0.009)	0.115 *** (0.014)	−0.036 (0.021)	−0.075 *** (0.005)	0.628 *** (0.039)	−0.149 *** (0.013)	−0.001 (0.008)
Region (reference: West)								
Midwest	0.382 *** (0.103)	−0.141 (0.270)	0.671 (0.352)	0.748 * (0.293)	0.074 (0.040)	1.807 * (0.886)	0.518 ** (0.186)	0.283 *** (0.077)
Northeast	0.004 (0.112)	1.142*** (0.293)	0.584 (0.382)	0.177 (0.318)	−0.063 (0.043)	1.257 (0.960)	0.034 (0.202)	0.016 (0.083)
South	0.136 (0.095)	0.004 (0.249)	1.601 *** (0.324)	1.430 *** (0.270)	0.071 (0.037)	2.707 *** (0.815)	0.782 *** (0.171)	0.300 *** (0.071)
Non-White (%):Typology	0.007 *** (0.0003)	0.003 *** (0.0003)	0.009 *** (0.0004)	0.004 *** (0.001)	0.002 *** (0.0001)	0.017 *** (0.001)	0.010 *** (0.0004)	0.003 *** (0.0002)
Log Likelihood	−94,621.5	−77,850.8	−107,450.8	−138,545.6	−38,074.71	−183,016.2	−106,300.1	−65,079.6
AIC	189,271	155,729.6	214,929.5	277,119.3	76,177.4	366,060.4	212,628.1	130,187.2
BIC	189,399.6	155,858.2	215,058.1	277,247.90	76,306	366,189	212,756.8	130,315.8

Coefficients represent the expected change in the crude prevalence rate (cases per 1000 adults) associated with a one-unit change in each predictor, holding all other variables constant. For the log-transformed total TRI variable, coefficients reflect the change in disease prevalence associated with a proportional change in toxic releases. Positive coefficients indicate higher disease prevalence associated with the predictor; negative coefficients indicate lower prevalence. Standard errors are shown in parentheses. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

In general, census tracts exposed to higher toxic releases are expected to have worse health outcomes. The coefficients for total toxic releases in Table 1 indicate that higher toxic releases correspond to a higher rate of cancer, COPD, diabetes, kidney diseases, arthritis, CHD, and stroke. Except for diabetes and kidney diseases, the associations were statistically significant. Contrary to expectations, the association between toxic releases and the crude asthma rate (Model 2) was negative, indicating that higher toxic releases are associated with lower asthma rates. We suspect that this result is due to the dataset including only health outcomes among adult residents, whereas asthma is more common in children and adults.

Across all models, higher toxic release from electric utilities and mining facilities was generally associated with worse population health outcomes at the census tract level. Statistically significant positive associations were observed for cancer, COPD, coronary heart disease, arthritis, and stroke. Although the estimated effect sizes were modest in absolute terms, these differences represent meaningful population-level shifts when aggregated across thousands of census tracts nationwide.

The strongest associations were observed for COPD and coronary heart disease, conditions with established links to inhaled pollutants. For example, a 10% increase in total toxic releases was associated with a measurable increase in COPD prevalence, consistent with existing evidence linking industrial emissions to respiratory morbidity. The negative observed association between toxic releases and adult asthma prevalence likely reflects limitations of the PLACES dataset, which excludes children, among whom asthma prevalence is highest. This finding should therefore be interpreted cautiously and does not imply a protective effect of pollution.

Socioeconomic and demographic covariates behaved largely as expected across models: higher poverty and lower educational attainment were consistently associated with worse health outcomes, while older median age strongly predicted higher chronic disease prevalence. Interaction analyses revealed that increases in the percentage of non-White residents were associated with disproportionately higher disease prevalence in rural areas compared with urban areas, highlighting important environmental justice dynamics.

4. Discussion

This study identified consistent associations between higher toxic releases from electric utilities and mining facilities and increased the prevalence of multiple chronic diseases at the census tract level. These findings align with proper research linking industrial emissions to respiratory, cardiovascular, and oncologic outcomes and extend this literature by focusing on two major emissions sectors (electric utilities and mining facilities) at the national scale with fine spatial resolution [1,7,21,22]. The observed associations are biologically plausible. Heavy metals such as arsenic, chromium, and nickel, commonly emitted by mining and utility operations, are known to include oxidative stress, inflammation, and genotoxic damage, mechanisms implicated in carcinogenesis and cardiovascular disease [15–17,24]. Combustion-related pollutants and VOCs contribute to airway inflammation and impaired lung function, providing a plausible pathway for the strong associations observed with COPD [25,26]. Several limitations warrant careful consideration. The ecological design precludes causal inference and is subject to ecological fallacy. Exposure misclassification is likely due to pollutant dispersion beyond census tract boundaries and the use of the total mass of releases rather than toxicity weighted metrics. Residual confounding may persist due to unmeasured individual-level risk factors, and the adult only nature of PLCES data limits interpretation for childhood conditions such as asthma. Together, these findings suggest that industrial emissions from electric utilities and mining facilities are associated with the higher prevalence of selected chronic diseases at the census tract level and contribute

to population level health disparities in the United States, particularly in rural and higher non-White communities.

Electric utilities, coal mining, and metal mining are included in the TRI. For 2020, 3463 electric utilities, 716 metal mining, and 94 coal mining facilities reported environmental releases to the TRI program [27]. These sectors are important contributors to environmental contamination because they emit heavy metals such as arsenic, chromium, cobalt, lead, and nickel, as well as organic compounds such as benzene and aldehydes [12,13]. Several of these pollutants are known carcinogens or are associated with respiratory, neurological, or cardiovascular impacts [15,17,18].

Beyond toxicological relevance, these sectors also represent a substantial share of total reported TRI releases. Using the 2020 TRI dataset linked to census tracts in this study, total releases across all sectors were approximately 3.09 billion pounds. Of this total, electric utilities accounted for roughly 226 million pounds (~7% of total TRI releases), metal mining accounted for approximately 1.37 billion pounds (~44% of total TRI releases), and coal mining accounted for about 3.5 million pounds (<1%). Combined, these two sectors (electric utilities and mining) accounted for about 52% of all TRI releases in the dataset for 2020. This demonstrates that the selected sectors are not marginal emitters but constitute a dominant share of reported industrial releases. The focus on these sectors is therefore justified on two grounds. First, they contribute a large fraction of TRI-reported environmental contamination, particularly through mining releases that dominate national totals. Second, the chemical profile of these sectors includes metals and combustion-related pollutants with well-established links to chronic diseases. Concentrating on these sectors allows the analysis to examine sources that are both quantitatively significant and biologically plausible contributors to population health risks.

At the same time, TRI does not represent the total emissions from all sources in the U.S. The program excludes many pollutants (e.g., most greenhouse gases and PM2.5) and only covers facilities that meet the reporting thresholds. Thus, the results should be interpreted as reflecting a large share of reported industrial point-source releases, and not the full universe of environmental exposures.

The associations between health outcomes and socio-economic factors examined in the models are mostly consistent with the literature. For each health outcome, higher poverty is associated with a higher crude disease rate across census tracts, all else being equal. The associations between the crude rates of diseases and poverty were statistically significant across all health outcomes examined. On average, census tracts with a higher percentage of female residents had a statistically significantly higher crude rate of disease for each health outcome. More educated census tracts, as measured by the percent of residents with a bachelor's degree or higher, were associated with lower crude rates across all health outcomes examined, while holding all other factors constant. As expected, older census tracts, as measured by median age, had worse health outcomes. Across all health outcomes but asthma, older median ages of census tracts corresponded to significantly higher crude rates of diseases. The association between median age and the crude rate of asthma was positive but statistically nonsignificant. This outcome is likely due to the dataset excluding children under 18 years old.

The models found significant associations between health outcomes and the racial composition of census tracts. The directions of associations, however, are inconsistent across health outcomes. In urban areas, a higher percentage of non-White population positively corresponded to crude rates of asthma, diabetes, kidney diseases, and stroke and negatively corresponded to rates of cancer, COPD, arthritis, and CHD. Furthermore, the associations between health outcomes and the percentage of non-White residents varied across geographies. Across all eight models, the coefficients of the interaction between

urban typology and percentage of non-White residents were positive and significant. This result indicates that as the share of non-White residents increases, rural census tracts have significantly higher increases in crude rates of diseases than their urban counterparts, all else being equal.

The models showed significant differences in health outcomes between urban and rural areas. Rural areas had significantly higher rates of asthma, COPD, and arthritis than urban areas. Meanwhile, rural areas had significantly lower rates of cancer, kidney diseases, and CHD than urban areas. The models found no statistically significant associations between urban/rural typologies, diabetes, and stroke.

Finally, health outcomes varied across census regions. The positive coefficients for the categorical census region variables indicate that on average, census tracts in the West region have better health outcomes, or lower rates of diseases than those in the other regions. Some of these differences across geographies were statistically significant. While census tracts in the Midwest had lower rates of asthma and census tracts in the Northeast had lower rates of kidney disease than census tracts in the West, the differences were not statistically significant.

From the environmental justice perspective, the findings indicate that rural communities with higher proportions of non-White residents experience steeper increases in disease prevalence associated with toxic releases than urban communities. This pattern likely reflects overlapping vulnerabilities related to facility siting, economic disadvantage, limited healthcare access, and reduced regulatory enforcement in rural areas [9,26].

Several features of TRI reporting and data structure over 2016–2020 could have influenced the certainty of multi-year comparisons, but they also strengthened the interpretability of the 2020 analysis used in this study. First, the TRI chemical list and reporting framework evolve over time as EPA adds chemicals or adjusts reporting requirements. Such changes can affect the total release amounts and chemical composition even if the underlying industrial activity remains stable. For example, newly listed substances or modifications to reporting thresholds can introduce discontinuities in multi-year totals. These shifts make long time-series comparisons based purely on total pounds more uncertain unless the chemical universe is harmonized across years. Second, the TRI totals reflect the mass of releases rather than the toxicity weighting. Changes in the relative mix of chemicals—for instance, differences between high-volume waste streams from mining and smaller-volume but more toxic combustion-related emissions from utilities—can influence the public-health relevance of releases over time without necessarily changing the total pounds. This compositional variability means that cross-year comparisons of total releases should be interpreted cautiously.

However, the year 2020 provides particularly valuable insights for this study for two reasons. First, the TRI reporting framework in 2020 is well-defined and comprehensive within the program's scope, meaning that sector comparisons and chemical profiles for that year are internally consistent. Second, and more importantly for this study's design, 2020 is the first year in which the CDC PLACES dataset provides nationwide tract-level health estimates rather than the more limited geographic coverage available in earlier years. This substantially improves the spatial alignment between pollution sources, demographics, and health outcomes.

Recommendations for Policy and Research

Based on the results of this study and several others, there are risks associated with exposure to chemicals released by electric utilities and metal and coal mining facilities in the U.S.. Although there has been a significant decrease in the rate of reported chemical emissions from 2016–2020, the timeline that was included in this study, we can still see a

positive correlation between these releases and chronic diseases. The federal regulatory bodies, including the EPA, should emphasize limiting toxic pollutants from all industries. In addition, the EPA should request data from sectors on greenhouse gas and PM2.5 emissions to fill data gaps. The NATA program is beneficial. However, it is limited to air and released only periodically. This program needs to be expanded to make it easy for users to add media and the release year. RSEI is a friendly tool, but it is limited. The interpretations of hazard scores are complex and difficult to understand.

The study is subject to limitations. First, the ecological design precludes inference about individual level risk (ecological fallacy). Second, exposure misclassification is likely because emissions can disperse across tract boundaries and because TRI mass does not reflect toxicity. Third, the cross-sectional nature of the data prevents the assessment of temporal relationships or causality. Fourth, the limitation of the PLACES dataset is the exclusion of the population under 18 years old, and finally, unmeasured level factors may confound associations.

5. Conclusions

This national ecological analysis demonstrates that census tracts with higher toxic releases from electric utilities and mining facilities experience a higher prevalence of several chronic diseases, including COPD, CHD, arthritis, and cancer. Although effect sizes were modest, the consistency of associations across multiple outcomes and the concentration of emissions in vulnerable communities highlight important environmental justice concerns. These findings support the need for stronger industrial emission controls, improved TRI reporting that incorporates toxicity and dispersion, and expanded health surveillance that includes children and longitudinal follow-up. Future research should integrate multi-year exposure data, toxicity weighted indices such as RSEI, and individual level health information to better characterize causal pathways and guide equitable public health interventions. These results underscore the need for targeted emissions reduction strategies and improved environmental surveillance in communities disproportionately burdened by industrial pollution.

Author Contributions: For research articles with multiple authors, a short paragraph outlining each author's individual contributions must be provided. Conceptualization, A.A. (Azita Amiri); methodology, A.A. (Azita Amiri), X.D. and M.F.; formal analysis, X.D. and S.Z.; investigation, A.A. (Armita Amiri); resources, A.A. (Azita Amiri); data curation, A.A. (Armita Amiri) and X.D.; writing—original draft preparation, A.A. (Azita Amiri) and S.Z.; writing—review and editing, M.F.; visualization: S.Z.; supervision, A.A. (Azita Amiri); project administration, A.A. (Azita Amiri) All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study is a retrospective study, and the main data used in the research are publicly available the EPA's Toxic Release Inventory (TRI) data and PLACES health data with no identifiers. The specific information can be found in Section 2.

Informed Consent Statement: Patient consent was waived because this research was a retrospective study.

Data Availability Statement: The data presented in this study are openly available in TRI Basic Data Files: Calendar Years 1987–Present, available online: https://www.epa.gov/toxics-release-inventory-tri-program/tri-basic-data-files-calendar-years-1987-present, accessed on 16 February 2026.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

TRI	Toxic Release Inventory
EPA	Environmetnal Protection Agency
PLACES	CDC PLACES health dataset
CDC	Centers for Disease Control and Prevention
COPD	Chronic Obstructive Pulmonary Disease
CHD	Coronary Heart Disease
VOCs	Volatile Organic Compounds
U.S.	United States
NATA	National Air Toxics Assessment
RESI	Risk-Screening Environmental Indicators

References

1. Gong, X.; Lin, Y.; Bell, M.L.; Zhan, F.B. Associations between maternal residential proximity to air emissions from industrial facilities and low birth weight in Texas, USA. *Environ. Int.* **2018**, *120*, 181–198. [CrossRef] [PubMed]
2. Moore, R.J.H.; Hotchkiss, J.L. The importance of toxicity in determining the impact of hazardous air pollutants on the respiratory health of children in Tennessee. *Environ. Pollut.* **2016**, *216*, 616–623. [CrossRef] [PubMed]
3. Annesi-Maesano, I.; Forastiere, F.; Balmes, J.; Garcia, E.; Harkema, J.; Holgate, S.; Kelly, F.; Khreis, H.; Hoffmann, B.; Maesano, C.N.; et al. The clear and persistent impact of air pollution on chronic respiratory diseases: A call for interventions. *Eur. Respir. J.* **2021**, *57*, 2002981. [CrossRef] [PubMed]
4. Zhang, Z.; Chai, P.; Wang, J.; Ye, Z.; Shen, P.; Lu, H.; Jin, M.; Gu, M.; Li, D.; Lin, H.; et al. Association of particulate matter air pollution and hospital visits for respiratory diseases: A time-series study from China. *Environ. Sci. Pollut. Res. Int.* **2019**, *26*, 12280–12287. [CrossRef] [PubMed]
5. Sakhvidi, M.J.Z.; Lequy, E.; Goldberg, M.; Jacquemin, B. Air pollution exposure and bladder, kidney and urinary tract cancer risk: A systematic review. *Environ. Pollut.* **2020**, *267*, 115328. [CrossRef] [PubMed]
6. Yang, B.Y.; Fan, S.; Thiering, E.; Seissler, J.; Nowak, D.; Dong, G.H.; Heinrich, J. Ambient air pollution and diabetes: A systematic review and meta-analysis. *Environ. Res.* **2020**, *180*, 108817. [CrossRef] [PubMed]
7. Kim, H.; Kim, W.H.; Kim, Y.Y.; Park, H.Y. Air Pollution and Central Nervous System Disease: A Review of the Impact of Fine Particulate Matter on Neurological Disorders. *Front. Public Health* **2020**, *8*, 575330. [CrossRef] [PubMed]
8. Johnson, R.; Ramsey-White, K.; Fuller, C.H. Socio-demographic Differences in Toxic Release Inventory Siting and Emissions in Metro Atlanta. *Int. J. Environ. Res. Public Health* **2016**, *13*, 747. [CrossRef]
9. Pastor, M., Jr.; Sadd, J.L.; Morello-Frosch, R. Waiting to Inhale: The Demographics of Toxic Air Release Facilities in 21st-Century California. *Soc. Sci. Q.* **2004**, *85*, 420–440. [CrossRef]
10. Neumann, C.M.; Forman, D.L.; Rothlein, J.E. Hazard screening of chemical releases and environmental equity analysis of populations proximate to toxic release inventory facilities in Oregon. *Environ. Health Perspect.* **1998**, *106*, 217–226. [CrossRef] [PubMed]
11. Wilson, S.M.; Fraser-Rahim, H.; Williams, E.; Zhang, H.; Rice, L.; Svendsen, E.; Abara, W. Assessment of the distribution of toxic release inventory facilities in metropolitan Charleston: An environmental justice case study. *Am. J. Public Health* **2012**, *102*, 1974–1980. [CrossRef] [PubMed]
12. EPA. Toxics Release Inventory (TRI) Program. Available online: https://www.epa.gov/toxics-release-inventory-tri-program/reporting-tri-facilities (accessed on 10 January 2026).
13. EPA. Hazard and Potential Risk of TRI Chemicals. Available online: https://www.epa.gov/toxics-release-inventory-tri-program/tri-chemical-hazard-information-profiles-tri-chip (accessed on 15 January 2026).
14. EPA. Understanding RSEI Results. Available online: https://www.epa.gov/rsei/understanding-rsei-results#what (accessed on 11 January 2026).
15. National Cancer Institute. Formaldehyde and Cancer Risk. NIH. Available online: https://www.cancer.gov/about-cancer/causes-prevention/risk/substances/formaldehyde/formaldehyde-fact-sheet (accessed on 11 June 2022).
16. National Cancer Institute. Arsenic. NIH. Available online: https://www.cancer.gov/about-cancer/causes-prevention/risk/substances/arsenic (accessed on 11 June 2022).
17. ATSDR. Lead-ToxFAQs. CDC. Available online: https://www.atsdr.cdc.gov/toxfaqs/tfacts13.pdf (accessed on 11 July 2022).
18. Centers for Disease Control and Prevention. Facts About Benzene. Available online: https://emergency.cdc.gov/agent/benzene/basics/facts.asp#: :text=The%20Department%20of%20Health%20and,of%20the%20blood%2Dforming%20organs (accessed on 11 July 2022).

19. EPA. 2014 NATA: Assessment Results. Available online: https://www.epa.gov/national-air-toxics-assessment/2014-nata-assessment-results (accessed on 11 January 2026).
20. EPA. 2017 AirToxScreen: Assessment Results. Available online: https://www.epa.gov/AirToxScreen/2017-airtoxscreen-assessment-results (accessed on 11 January 2026).
21. Hanchette, C.; Zhang, C.H.; Schwartz, G.G. Ovarian Cancer Incidence in the U.S. and Toxic Emissions from Pulp and Paper Plants: A Geospatial Analysis. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1619. [CrossRef]
22. Fortunato, L.; Abellan, J.J.; Beale, L.; LeFevre, S.; Richardson, S. Spatio-temporal patterns of bladder cancer incidence in Utah (1973–2004) and their association with the presence of toxic release inventory sites. *Int. J. Health Geogr.* **2011**, *10*, 16. [CrossRef]
23. Persico, C.L.; Johnson, K.R. The effects of increased pollution on COVID-19 cases and deaths. *J. Environ. Econ. Manag.* **2021**, *107*, 102431. [CrossRef] [PubMed]
24. Lv, J.; Li, X.; Shen, Y.; You, J.; Wen, M.; Wang, J.; Yang, X. Assessing Volatile Organic Compounds Exposure and Chronic Obstructive Pulmonary Diseases in US Adults. *Front. Public Health* **2023**, *11*, 1210136. [CrossRef] [PubMed]
25. Neofytou, E.; Tzortzaki, E.G.; Chatziantoniou, A.; Siafakas, N.M. DNA damage due to oxidative stress in chronic obstructive pulmonary disease (COPD). *Int. J. Mol. Sci.* **2012**, *13*, 16853–16864. [CrossRef] [PubMed]
26. Mohai, P.; Saha, R. Racial Inequality in the Distribution of Hazardous Waste: A National-Level Reassessment. *Soc. Probl.* **2007**, *54*, 343–370. [CrossRef]
27. EPA. TRI Basic Data Files: Calendar Years 1987-Present. Available online: https://www.epa.gov/toxics-release-inventory-tri-program/tri-basic-data-files-calendar-years-1987-present (accessed on 12 December 2025).
28. CDC. PLACES: County Data (Census Friendly Format), 2016 Release. Available online: https://chronicdata.cdc.gov/browse?q=PLACES%3A%20County%20Data%20(census%20Friendly%20Format)%2C%202016%20release&sortBy=relevance (accessed on 11 July 2022).
29. CDC. PLACES: Local Data for Better Health. Available online: https://www.cdc.gov/places/index.html (accessed on 11 July 2022).