**ENVIRONMENTAL HEALTH DISPARITIES ASSOCIATED TO AIR STRESSORS IN ALLEGHENY COUNTY, PENNSYLVANIA**

by

**Varun Patel**

BS – West Chester University of Pennsylvania, 2015

BSc – Maharaja Sayajirao University of Baroda, India, 2010

Submitted to the Graduate Faculty of

Environmental and Occupational Health

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This essay is submitted

by

**Varun Patel**

on

May 6, 2019

and approved by

**Essay Advisor:**

James Peterson, PhD

Associate Professor

Environmental and Occupational Health

Graduate School of Public Health

University of Pittsburgh

**Essay Readers:**

Shaina Stacy, PhD, MPH

Post-Doctoral Scholar

Epidemiology

Graduate School of Public Health

University of Pittsburgh

Jian-Min Yuan, MD, PhD

Professor

Epidemiology

Graduate School of Public Health

University of Pittsburgh

Copyright © by Varun Patel

2019

Jim Peterson, Ph.D.

**ENVIRONMENTAL HEALTH DISPARITIES ASSOCIATED TO AIR STRESSORS IN ALLEGHENY COUNTY, PENNSYLVANIA**

Varun Patel, MPH

University of Pittsburgh, 2019

**ABSTRACT**

Allegheny County, located in southwestern Pennsylvania, has a rich history of industry that includes glass making, steel production, coal-fired power plants, and mining-associated activities such as coke processing. Facilities related to these industries contribute significantly to air pollution, releasing pollutants such as butadiene, formaldehyde, and acetaldehyde into the air. Coke oven emissions are major air stressors in southeast regions of the county. Coke oven emissions are predominantly released from large ovens used in heating coal to produce coke in steel and iron manufacturing facilities. The emissions are complex mixtures of dust, vapors, and gases that typically include carcinogens such as cadmium and arsenic. In addition, traffic-related pollutants including diesel particulate matter also contribute to poor air quality. Spatial associations between cancer incidence and mortality with air pollution are well studied in several cities in the United States and around the world. However, this study is an attempt to examine the association at a smaller scale i.e., census tract level of Allegheny County. For this study, we used United States Census data, the Pennsylvania Cancer Registry, and the Environmental Protection Agency’s (EPA) National Air Toxic Release Assessment (NATA) data for geospatial analysis at the census tract level for Allegheny County. Spatial analysis was used to investigate the association between ambient concentrations of air toxics, lung cancer incidence (*N* = 6,435) and socioeconomic status (SES) (race/ethnicity and income) in the county from 2010-2015. ArcGIS and QGIS were used to create interactive maps, and GeoDa was used to examine spatial and statistical relationships. We used global and local measures of spatial autocorrelation (Moran’s *I*) to identify clusters of tracts where lung cancer incidence was significantly higher. We identified a few local “hotspots” of higher cancer incidence. We also found SES positively related to lung cancer incidence as well as ambient levels of certain air toxics. This study revealed associations between lung cancer risk and environmental exposures and identified vulnerable communities where future resources could be allocated to help reduce the disproportionate public health burden.

TABLE OF CONTENTS

Introduction……………………………………………………………………………...01

Pollutants

1,3-Butadiene………………………….………………………………………………...04

**Acetaldehyde and Formaldehyde……….…………………………………………...…04**

**Benzene………………………………………………………………………….………06**

**Chloroform……………………………………………………………………….….….06**

**Coke Oven Emissions.…………………….…………………………………...………..07**

**Diesel PM………………………………………………………………………………..08**

**Ethylbenzene……………………………………………………………………………09**

**Trichlorethylene……………………………………………………………...................09**

**MATERIALS AND METHODS**

**Study Population……………………………………..……………………..…………..11**

**Cancer Incidences…………………………………………………………….………...11**

**Environmental Exposure…………………………….…………..……….….…………12**

**Toxic Release Inventory……………………………….…………………….………….13**

**Environmental Justice………………………..……….…………………….………….13**

**Geospatial and Statistical Analysis…………………….……………………………....14**

**RESULTS……………………………………………………………………………………….17**

**CONCLUSION……………………………………………………………………...………….25**

**BIBLIOGRAPHY………………………………………………………………………………27**

List of tables

Table 1: Distribution of air pollution concentration (µg/m3) in Allegheny County……...……...12

Table 2: Descriptive statistics for cancer incidence rates by race, gender, age and SESindex…....20

Table 3: Association of lung cancer incidence rate adjusted with SES…………………………...21

Table 4: Association of exposure to pollutants and lung cancer variables adjusted for SES……...23

Table 5: Association of lung cancer and SES, adjusted for Coke Oven Emissions and TCE…......24

List of figures

Figure 1: Moran’s *I*…………………………………………………………….…………………15

Figure 2: Toxic Release Sites and EJ Tracts in Allegheny County……………...………………...17

Figure 3: Lung Cancer Rates in African Americans in Allegheny County (2010 – 2015)………...18

Figure 4: Local Clusters for Each Pollutant………………………………………………….…...19

Figure 5: Local Clusters for Total Lung Cancer Incidence Rates……………………………...….20

# Introduction

Cancer is the second leading cause of deaths, which influences a significant public health problem in the United States and across the world. In 2015, the total number of cancer related deaths were 595,930 (22%) out of 2,712,630 deaths and 598,038 (22%) out of 2,744,248 deaths in 2016 in the United States. In 2019, The American Cancer Society estimates 1,762,450 new cancer cases and 606,880 cancer-related deaths in the United States (Siegel, Miller, & Jemal, 2019a). Similarly, cancer is also a leading cause of death in Pennsylvania. In 2011, over 77,000 people had a diagnosis of invasive cancer and approximately 28,500 died of cancer (Pennsylvania Department of Health, 2015). A total 79,890 new cases for selected cancers[[1]](#footnote-1) are projected in 2019 in Pennsylvania. ﻿Beyond new diagnoses each year and the number of lives lost, cancer also inflicts a great financial and emotional burden on cancer patients and their loved ones. The estimated national expenditure for cancer care in 2017 was $147.3 billion in the United States (National Institutes of Health, 2018). In 2010, the estimated cancer cost was $7.3 billion in Pennsylvania.

Ambient air pollution has various effects on human health including cancer, cardiovascular diseases, and other non-cancerous acute and chronic diseases (Linder, Marko, & Ken, 2008). More than 30% of cancers may be preventable by lowering ambient air pollution exposure levels, changing lifestyle factors, and with advanced screening technologies to detect cancer at early stages (Pennsylvania Department of Health, 2015).

﻿Socioeconomic factors such as poverty, inadequate education, lack of health insurance, and location of residence are equally important as biological factors of cancer (Ward et al., 2004). It is well documented that disadvantaged populations are more likely to have cancer than affluent populations (cite). ﻿For all cancer sites combined, residents of poorer areas (those with greater than or equal to 20% of the population below the poverty line) have 13% higher death rates from cancer in men and 3% higher rates in women compared with more affluent areas (less than 10% below the poverty line) (Ward et al., 2004). The association of specific cancer prevalence and cancer mortality varies at individual or area levels. ﻿The major behavioral determinants of cancer (e.g., smoking, diet, alcohol use, obesity, physical inactivity, reproductive behavior), occupational and environmental exposures, and cancer screening are themselves substantially influenced by individual and area-based socioeconomic factors (Singh, Williams, Siahpush, & Mulhollen, 2011). Poor families tend to live close to industrial areas because they cannot afford to live in the suburbs. It is also likely that some of these family members are employed at manufacturing facilities. Thus, their exposure is significantly higher than those living far from these facilities. Poor housing also plays a major role in terms of exposure to carcinogens such as asbestoses and heavy metals such as lead (Boffetta & Nyberg, 2003).

Air pollution is a complex mixture of particulate matter and gaseous compounds. Outdoor air pollution consists of particulate matter (PM), nitrogen oxides (NOx), ozone, sulfur oxides (SOx), diesel exhaust, carbon dioxide (NO2), carbon monoxide (NO) and coke oven emissions. Some pollutants such as acetaldehyde, benzene, butadiene, formaldehyde, and trichloroethylene are present at various levels in the environment. These pollutants predominantly release from refineries, steel manufacturing facilities, power plants, glass manufacturing factories and fuel combustion in automobiles (Hemminki & Pershagen, 1994; Lewtas, 2007; Wilhelm & Ritz, 2005).

Three major factors are responsible for many cancers: 1) DNA mutations, which initiate during stem cell replication or are induced by environmental factors, 2) hereditary, and 3) unknown (or a combination of both) factors (Tomasetti & Vogelstein, 2015). Environmental factors include physical and chemical, behavioral, and social components. Physical and chemical components such as air pollution from various industries, automobile exhaust, water, and food contamination; behavioral components such as smoking, unhealthy diet and lack of exercise; and social components such as poverty, occupation, and lack of education could play a huge role in cancer epidemiology. The aim of this paper is to discuss the environmental risk factors of lung cancer accounting for socioeconomic factors. Ultrafine particles (UFP) of engine exhaust or diesel PM has attracted a great deal of focus because of toxicological associations to morbidity and mortality of chronic diseases including cancer (Parent, Rousseau, Boffetta, Cohen, & Siemiatycki, 2007; Silverman, 2017). Pollutants such as particulate matter, diesel and coke oven emissions can generate reactive oxygen species (ROS) that increase risk due to their toxicological mechanism to cause oxidative stress in the body. ﻿Oxidative stress is typically assessed as elevated levels of oxidized biomolecules, e.g. oxidative DNA damage, which is relevant for carcinogenesis (Avogbe et al., 2005). It has been reported that toxic effects of ROS on human cells may end in oxidative injury leading to programmed cell death i.e. apoptosis (Uttara, Singh, Zamboni, & Mahajan, 2009).

The complexity of air pollution often makes it harder to determine one or two pollutants as causative factors for a specific health problem. Prior epidemiological studies at the county level have established an association between air pollution and different kinds of cancers, mainly lung cancer (Martien, Ph, Lau, Tanrikulu, & Ph, 2014; Pope III et al., 2002). However, air toxic concentrations vary at spatial levels much smaller than a county. For this epidemiological study, we examined the association of nine pollutants we hypothesize have a spatial relationship with lung cancer prevalence using cancer registry data from 2010 to 2015.

## Pollutants

**1,3-butadiene:** The main use of 1,3-butadiene is in manufacturing rubbers and plastics. It is also used in acrylic polymers manufacturing. In addition, automobile exhaust, cigarette smoke and other combustion such as forest fires are major sources of 1,3-butadiene (United States Environmental Protection Agency, 2009). The International Agency for Research on Cancer (IARC) has identified 1,3-butadiene as a Group 1 carcinogen to humans. Several epidemiological studies have suggested associations of butadiene with leukemia and ﻿non-Hodgkin lymphoma in workers of rubber and plastic manufacturing (Sathiakumar, Brill, & Delzell, 2009) and with lung cancer (Sharma et al., 2019). Animal studies have identified butadiene ﻿as a leading factor for early induction and significantly increased incidence of hemangiosarcoma of the heart, malignant lymphomas, alveolar-bronchiolar neoplasms, and squamous cell neoplasms in rats (Miller, 1984). In the 2005 NATA assessment, the exposure levels of butadiene in Allegheny County recorded were relatively lower (0.09 ppm to 0.20 ppm) than the Occupational Safety and Health Administration’s (OSHA) permissible exposure level (PEL) – 1 ppm and short-term exposure level (STEL) – 5 ppm (United States Environmental Protection Agency, 2009).

**Acetaldehyde and Formaldehyde:** Acetaldehyde and formaldehyde are the most common aldehydes found in ambient air. The main use of acetaldehyde is as an intermediate in the synthesis of other chemicals. It is used in perfumes, polyester resins, and basic dyes production. Acetaldehyde is also used for preservatives, rubber, and tanning industries (United States Environmental Protection Agency, 2016a). It is a natural product of photo-oxidation and combustion of hydrocarbons in the atmosphere. In a German survey study, nine cancer cases were identified in workers, where the main process was dimerization of acetaldehyde in an aldehyde factory. Formaldehyde is mainly used in many household products and building materials such as pressed-wood products – particleboard, plywood, fiberboard, glues, adhesives, permanent-press fabrics, paper product coatings, and certain insulation materials. Both acetaldehyde and formaldehyde break down rapidly in the air and water and dilute within hours. Similar to acetaldehyde, exposure of formaldehyde was linked with cancer in laboratory animals. Nasal cavity cancers and leukemias are the most common cancers observed in rats after inhaling formaldehyde (American Cancer Society, 2014). There have been adverse effects documented due to formaldehyde exposure from polluted air in countries like China, Egypt, Indonesia, and parts of the United States. Formaldehyde is widely used in hair straightening and other cosmetic products. Hairstylists are highly exposed to formaldehyde from hair products in beauty salons (Aglan & Mansour, 2018). Acute toxicity of formaldehyde shows symptoms like ﻿irritated eyes, tearing, sneezing, coughing, chest congestion, fever, heartburn, lethargy, and loss of appetite. Chronic exposure of formaldehyde may lead to long-term health effects including neurotoxicity, ﻿pulmonary function damage, ﻿hematotoxicity, ﻿reproductive toxicity, and allergic asthma (Tang et al., 2009). In a case study, a previously healthy woman was diagnosed with ﻿pancytopenia – anemia just after moving to another apartment. The formaldehyde concentration in the air was found to be four times higher than the national concentration (0.1 µg/m3) in the new apartment (Huang, Zou, & Deng, 2007). Indoor exposure for both acetaldehyde and formaldehyde have a higher magnitude of health effects than the outdoor (Salariino et al., 1985; Samet, Marbury, & Spengler, 1987).

**Benzene:** Benzene is a well-studied pollutant in terms of health effects compared to other pollutants. Benzene has been associated with leukemias such as acute non-lymphocytic leukemia (ANLL) and acute myeloid leukemia (AML) in refinery workers. It is used mainly in making plastics, rubber, lubricants, drugs, detergent, and pesticides. About four decades ago, it was predominantly used as an industrial solvent. Nowadays, the main source of benzene is from automobile exhaust as it is a natural component of gasoline. People may get exposed to benzene through emission from industries such as paint and chemical processing factories. Heavily trafficked and surrounding areas of gas stations also have a high level of benzene in the air. Different levels of benzene are present in the ambient air in both urban and rural areas, with urban air typically having higher levels of benzene. ﻿Myelogenous leukemia, bone marrow depression, and blood cell-related cancers have been linked to high dose exposure of benzene in rodents and humans (American Cancer Society, 2016; Avogbe et al., 2005). ROS produced when benzene goes under ﻿hepatic metabolism, generating hydroquinone, phenol and other compounds with the ability of redox cycling can cause DNA damage. ﻿The immune-toxic effects of benzene exposure in acute and chronic cases have been reported in both experimental and epidemiologic studies. The published epidemiologic studies, regarding the toxic effects of benzene on immune function, are related to occupational exposure. It is still unclear whether present benzene levels in our environment play any role in the incidence of immunological problems (Bahadar, Mostafalou, & Abdollahi, 2014).

**Chloroform:** Chloroform is mainly used to make other chemicals. It releases in the air as a result of its formation in the chlorination of drinking water, wastewater and swimming pools. Chloroform enters the environment from pulp and paper mills, hazardous waste sites, and sanitary landfills emissions (United States Environmental Protection Agency, 2000). In addition, human exposure of chloroform also occurs through drinking water and can absorb through the skin during bath/shower. There is not enough evidence available to link chloroform with carcinogenic effects. However, it can be detected in blood, urine, and body tissues. Most inhalation exposure data gathered from clinical settings, where chloroform was used in anesthesia. Animal studies have shown signs in kidney and liver damage in rats and limited information available on reproductive or developmental effects due to chloroform exposure (Agency for Toxic Substances and Disease Registry (CDC), 1997).

**Coke Oven Emissions (COE):** Coke oven emissions largely release from large ovens used in heating coal to produce coke in steel and iron manufacturing facilities. The emissions are complex mixtures of dust, vapors, and gases that typically include carcinogens such as cadmium, arsenic, 60 other organic compounds and 40 polycyclic aromatic hydrocarbons (PAHs). These PAHs are released from ﻿various industrial stacks, furnace, basic oxygen furnace, coke oven, electric arc furnace, heavy oil plants, and power plants. ﻿Several PAHs are known to be mutagenic and carcinogenic toward rodents in the laboratory, and potential carcinogens to humans (Yang, Lee, Chen, & Lai, 1998). There is epidemiologic evidence available suggesting that COE rich in polycyclic aromatic hydrocarbons can cause lung cancer in humans (Cohen, 2000; Tornqvist & Ehrenberg, 1994). Similar to benzene, most COE studies have been conducted on occupational cohorts – steelworkers. COE was linked to employment in coke production and cancer of the skin, urinary bladder, and respiratory tract before 1950. Thereafter, numerous studies conducted in different countries linked exposure to COE with lung cancer. Other studies have reported an increased number of hematopoietic and renal cancers in coke workers in the United States (National Toxicology Program (NIH), 2016). IARC Monographs has evaluated COE as a causative factor for lung cancer in 1984; since then, no new studies have been published.

**Diesel PM:** Air pollution, containing a high level of UFP (2.5 – 10 µm) is a leading public and environmental health issue in many cities around the world. Motorized traffic is the major source of particulate matter, especially diesel particulates. Diesel PM releases from heavy-duty vehicles such as buses and trucks. In cities like Pittsburgh, where the public transportation is heavily depended on diesel engine buses, the exposure from bus exhaust is higher for people who are living in the city and close to the busy busways. Several studies have suggested an association exists between living close to heavily trafficked roadways and adverse health effects in the respiratory system (Janssen, Vliet, Harssema, & Brunekreef, 2001).

Diesel PM includes a combination of arsenic, benzene, formaldehyde, and nickel. Thus, it can potentially contribute to mutations that can lead to cancer. Particulate matter can penetrate into respiratory airways and ﻿deposit in the respiratory bronchioles and alveoli. ﻿The most important factors present are transition metals with redox properties, persistent free radicals, redox-cycling of quinones, polycyclic aromatic hydrocarbons (PAHs) and volatile organic compounds (VOCs) which may be metabolically activated to ROS that can react to form bulky adducts or strand breaks on cellular DNA. The association of oxidative stress and ﻿incidence of malignant respiratory diseases due to inflammation, activation of transcriptional factors and DNA damage are well studied in cellular biology (Valavanidis, Vlachogianni, Fiotakis, & Loridas, 2013). According to the California Environmental Protection Agency (CAEPA), long term exposure of diesel PM has the highest cancer risk compared to any other pollutants present in the air (Office of Environmental Health Hazard Assessment, 2001). An Eastern European study examined reproductive effects due to PM exposure and found elevations of reproductive effects in male (lower sperm count) and females (lower birthweight and embryo toxicity) who were living in industrial areas. They also suggested a link between exposure of particulate matter and ﻿intrauterine growth retardation (IUGR) or fetal growth retardation leading to low birth weights (Lewtas, 2007). Another case-control study found that risk of low birth weight ﻿increased with exposure to ambient air pollution related to the petroleum combustion products emitted from vehicles in Southern California (Lewtas, 2007).

**Ethylbenzene:** Ethylbenzene, naturally found in oil, is predominantly used in making styrene in the United States. It is also used in fuels and solvents such as paints, carpet glues, gasoline, paint, and automotive products. Ethylbenzene is present in low levels in rural areas and slightly higher in urban areas, especially close to tunnels and gas stations. Cigarette smoke also has been identified as a source of exposure to this chemical. People who are living near manufacturing facilities, petroleum refineries, and hazardous waste disposal sites and those working or residing in high traffic areas may potentially be exposed to ethylbenzene (Martenies et al., 2017). Very limited data are available on ethylbenzene’s toxicity and/or carcinogenicity in humans. However, several studies have suggested systematic effects of ethylbenzene in animals. Acute and intermediate exposure was associated with respiratory irritation, changes to the liver (increased organ weights and induction of microsomal enzymes), and effects on the hematological system (decreased platelets and increased leukocyte counts). Chronic exposure is associated with adverse effects to the liver (necrosis and hypertrophy), kidney (nephropathy and hyperplasia), and endocrine system (thyroid and pituitary hyperplasia) in animals (Agency for Toxic Substances and Disease Registry (CDC), 2010). There is not enough evidence to associate ethylbenzene with reproductive effects (United States Environmental Protection Agency, 2016b).

**Trichloroethylene (TCE):** Trichloroethylene is a colorless and toxic volatile organic (VOC) compound. The primary use of TCE is as a solvent to make refrigerant chemicals. It is also used as metal degreasing and dry cleaning agent (United States Environmental Protection Agency, 2017). TCE enters the environment from manufacturing and processing facilities. It breaks down slowly in the air, thus, remains in the air for a long period of time. People are exposed to TCE by inhaling, as well as through drinking contaminated water. The levels of TCE in ambient air have been declining recently. Based on numerous human studies showing a causal association between exposure to TCE and increased lung cancer risk, as well as similar findings in animals, TCE has been upgraded to known to be a human carcinogen (Agency for Toxic Substances and Disease Registry (CDC), 2016). Several human studies have associated TCE with kidney cancer and non-Hodgkin lymphoma. An animal study has suggested plausible evidence for the development of the morphological lesion and biochemical changes in the mouse lung after exposure to trichloroethylene (Odum, Foster, & Green, 1992). A systematic review study was conducted to identify the evidence of an association between TCE exposure and kidney cancer, non-Hodgkin lymphoma and liver cancers. After reviewing twenty-four studies they conclude that (occupational) exposure of TCE increases kidney cancer 2-fold.

The objective of this study is to conduct geospatial and statistical analysis to examine associations between environmental exposures (to the agents listed above) and risk of lung cancer at a small geographic level, i.e. the census tract. For this study, we mapped disparities in lung cancer for each census tract and focused only on ambient exposure levels of air toxics.

## Materials and Methods

**Study Population:** Allegheny County is our study population for this study. The county consists of 416 census tracts[[2]](#footnote-2); the maps were created based on 2000 American Community Survey (ACS) census data for total population, population by gender, race and age groups. We also derived socioeconomic status (SES) data for poverty and educational attainment from ACS. The SES variable was based on total percentage of poverty by census tract. Similarly, the percentage of the population who at least have a high school education. We excluded the population under 40 years of age from this study, considering the lifetime risk of diagnosing with cancer increases after 40 years of age in both men and women (White et al., 2014). We only accounted for African American and Whites population in the study.

**Cancer Incidence:** For this study, we collected lung cancer (*N* = 6,348) data from the Pennsylvania Cancer Registry (PCR) from 2010 to 2015 for Allegheny County. Information for each cancer case was available including, ﻿International Classification of Diseases for Oncology (ICD-O-3) codes for primary sites, patient ID, race, sex, diagnosed date, last contacted and geographic coordinates. Recurring cancer incidence accounted only once by patient ID when patients were diagnosed for the first time, duplicate records were excluded. ﻿If a patient developed a second primary cancer, he or she was counted only once for total cancer and once for each specific cancer. ﻿Age-adjusted incidence rates expressed per 1,000 population were computed using the 2000 US standard population.

**Environmental Exposure:** We obtained ambient air pollution concentration levels for the nine pollutants described in the introduction – 1,3-butadiene, acetaldehyde, benzene, chloroform, COE, diesel PM, ethylbenzene, formaldehyde, and TCE, from EPA’s 2005 NATA program. NATA is a wide-ranging assessment of different pollutants in the United States, which estimates ambient air concentration at county and census tract level. About 180 different kinds of pollutants (plus diesel PM) are included in the 2005 NATA assessment (United States Environmental Protection Agency, 2018). These nine pollutants were selected for this study for two reasons; first, literatures have suggested direct or indirect links to lung cancer risk and second, their exposure level variability was much larger than other pollutants (Table 1). Although, benzene and ethylbenzene are possible carcinogenic to humans, there have not been enough studies conducted to link them with lung cancer.

Table 1: Distribution of air pollution concentration (µg/m3) in Allegheny County a

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean | Median | S.D.b | Range |
| 1,3-Butadiene | 0.110 | 0.106 | 0.014 | 0.093 – 0.195 |
| Acetaldehyde | 1.734 | 1.716 | 0.132 | 1.456 – 2.167 |
| Benzene | 1.356 | 1.296 | 0.320 | 0.916 – 3.789 |
| Chloroform | 0.107 | 0.102 | 0.023 | 0.071 – 0.364 |
| Coke Oven Emission | 0.018 | 0.012 | 0.023 | 0.004 – 0.257 |
| Diesel PM | 0.798 | 0.642 | 0.695 | 0.184 – 8.298 |
| Ethylbenzene | 0.208 | 0.188 | 0.107 | 0.058 – 0.939 |
| Formaldehyde | 1.771 | 1.741 | 0.221 | 1.333 – 2.728 |
| Trichloroethylene | 0.260 | 0.248 | 0.045 | 0.203 – 0.494 |

a Total 416 census tracts

b Standard Deviation

Data Derived from EPA 2005 NATA Program

**Toxic Release Inventory (TRI):** Almost all the pollutants described above are released from TRI sites in Allegheny County (except diesel PM). EPA developed a TRI program to track the management of certain chemicals that may pose a threat to human health and the environment. Manufacturing and processing facilities are required to report annually the quantity of chemical released (into air, water or landfill) to the environment and/or managed through recycling, energy recovery, and treatment (United States Environmental Protection Agency, 2019b). A total of 67 TRI sites are located in Allegheny County, including coke processing, steel works, paper mills, glass making, power plants, landfill sites, and wastewater treatment facilities. In order to create maps of TRI sites, we retrieved publicly available shapefiles from The Western Pennsylvania Regional Data Center (Western Pennsylvania Regional Data Center, 2019).

**Environmental Justice (EJ) Areas:** Environmental justice movement started in the early 1980s. This social movement focused on the fair distribution of environmental resources and burden in communities. The EPA defines EJ as:

*“Environmental justice is the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income, with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies. This goal will be achieved when everyone enjoys the same degree of protection from environmental and health hazards, and equal access to the decision-making process to have a healthy environment in which to live, learn, and work.”* (United States Environmental Protection Agency, 2019a).

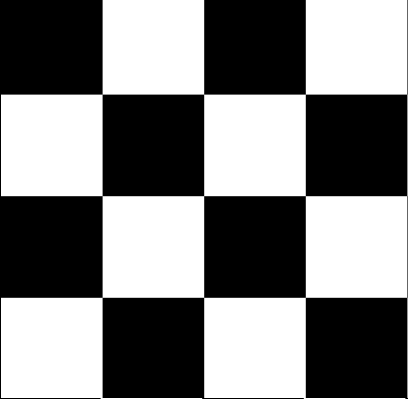
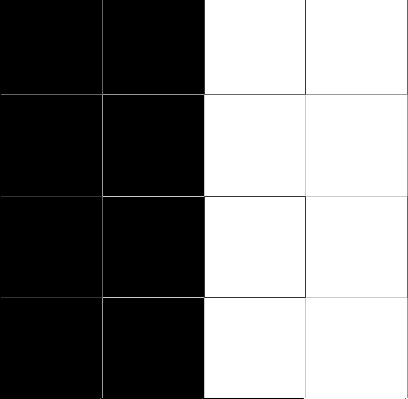
Pennsylvania Department of Environmental Protection (PA DEP) defined EJ tracts as any census tract where 30% or more population is a minority (identified as non-white) and at least 20% of the population lives below poverty level (Pennsylvania Department of Health, 2019). To examine the locations of TRI sites in proximity of EJ areas – census tracts, we downloaded the shapefiles for EJ tracts from Western Pennsylvania Regional Data Center (Western Pennsylvania Regional Data Center, 2016). We briefly examined the impact of air pollution for the EJ tracts because environmental exposure is higher in the communities with lower SES, and minorities have a higher health-related burden. Often times, they are not informed regarding environmental hazard present in their communities (Fabisiak, Jackson, Brink, & Presto, 2018).

**Geospatial and Statistical Analysis:** For this study, we utilized GeoDa software to conduct exploratory spatial data analysis. Spatial autocorrelation for the variables was explored using the global (test for clustering) and local (test for the cluster) Moran’s *I* statistics. The term “global” refers to testing for spatial autocorrelation for the entire study area at once and deriving a single value indicating whether spatial autocorrelation exists and if so, then at what strength. “Local” refers to testing for spatial autocorrelation with neighbors (Helbich, Leitner, & Kapusta, 2012). Census tracts which share common boundaries[[3]](#footnote-3) are called “neighbors”. Consequently, we used Moran’s *I* to evaluate the correlation among neighboring tracts and categorized spatial clustering strengths and levels. The values of Moran’s *I* may range from -1 to +1. Moran’s *I* = -1 indicates the lack of spatial correlation, in other words, variables are not clustered (Figure 1-A). Moran’s *I* = +1 when clusters are present (Figure 1-B), and Moran’s *I* = 0 represents perfect randomness (Figure 1-C) (Anselin, Syabri, & Kho, 2006; Sharma et al., 2019). We recalculated the Moran’s *I* by 999 permutations. The approach assesses the sensitivity of the results due to multiple comparisons (Anselin et al., 2006).

Global Moran’s *I* is important to examine the strengths of clustering; however, it does not provide any information about cluster locations. It also does not provide any information on what type of spatial correlation exists between clusters, for example, the correlation between higher values or lower values. In that case, local indicators of spatial association (LISA) specifies the measure of association for each spatial unit – census tract and provides information about types of spatial correlation (﻿clustering of similar values around that observation) (Anselin, 1995). We created LISA cluster maps for all pollutants and cancer rate to ﻿classify the locations by type of association. ﻿ High positive values refer to “hot spots” and high negative values to “cold spots”. As such, “hot spots” can be described as areas where tracts with high levels are surrounded by other tracts with high levels. In contrast, “cold spots” are clusters with low-level tracts surrounded by other low-level tracts. In addition to LISA maps, the standard output of a LISA analysis includes a Moran scatter plot depicting the distribution of the local statistic.

(B)

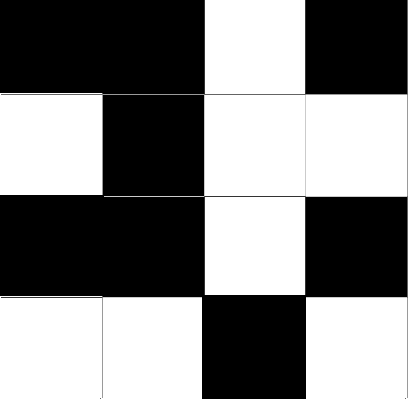
(A)

(C)

Moran’s *I* = +1

Moran’s *I* = -1



Moran’s *I* = 0

Figure 1: Moran’s *I*

Cancer rate variables were classified in the following classes: race (African Americans and Whites), gender (Males and Females), combination of race and gender (White Males, White Females, African American Males and African American Females) and age groups (40-49, 50-59, 60-69, 70-79, and 80 and over years). SPSS and GeoDa were utilized to performed simple linear regression to examine the association between lung cancer rates, the concentration of pollutants and SES by census tract. Since the sample size was smaller for some tracts, the data was not normally distributed within the county. Consequently, the log transformation was performed for cancer rate variables.

# Results

Out of a total 67 TRI site, 30 sites (45%) were located in EJ tracts, and 16 (24%) located close proximity to EJ areas (Figure 2). Almost half of the TRI sites were also located close to Pittsburgh’s three major rivers – Allegheny, Monongahela and Ohio rivers. Recently an environmental group reported about 40% of industrial facilities in Allegheny County have exceeded their pollution limit at least once in 2017 (www.alleghenyfront.org) (Lancianese, 2018).

A close up of a map

Description automatically generated

Figure 2: Toxic Release Sites and EJ Tracts in Allegheny County

To analyze the impact of lung cancer in the African American population, a map was created to demonstrate cases of lung cancer by census tract (Figure 3). About 83% of African American lung cancer cases were located in EJ areas. EJ tracts represent both percentages of poverty and minority. Figure 3 suggests that there is higher incidence of lung cancer in poor areas with higher minorities.

A picture containing text, map

Description automatically generated

Figure 3: Lung Cancer Rates in African Americans in Allegheny County (2010 – 2015)

The most impacted areas with air pollution are in the city limits of Pittsburgh. However, the areas with high coke oven emissions are clustered significantly higher in the southeast side of Allegheny County, which indicates the lower air quality due to coke works facilities (Figure 4). The higher exposure for all pollutants showed spatial clustering in the city of Pittsburgh neighborhoods, with Moran’s *I* ranging from 0.20 to 0.74. Diesel PM, formaldehyde and acetaldehyde are clustered in the trafficked areas in Downtown. In addition, the impacted areas are relatively less affluent than the less impacted areas. Similarly, the total lung cancer incidence rates significantly clustered (Moran’s *I* = 0.66) in the same areas where air pollution is relatively higher (Figure 5).

A total 6,348 lung cancer cases were observed in Allegheny County from 2010 – 2015. There were large variations in the rates of cancer cases. The mean of the total lung cancer incidence rate is 10.74 cases per 1,000 population. The variability was much larger in African American males and females. The increased cancer incidence rate was observed with increments of age (Table 2).

(A) 1,3-Butadiene (B) Acetaldehyde (C) Benzene

A close up of text on a white background

Description automatically generatedA close up of a logo

Description automatically generatedA close up of a logo

Description automatically generated

Moran’s *I* = 0.70 Moran’s *I* = 0.74 Moran’s *I* = 0.65

(D) Chloroform (E) COE (F) Diesel PM

A close up of a logo

Description automatically generatedA close up of a logo

Description automatically generatedA close up of a logo

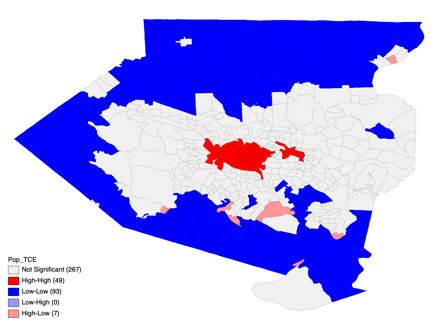
Description automatically generated

Moran’s *I* = 0.49 Moran’s *I* = 0.74 Moran’s *I* = 0.20

(G) Ethylbenzene (H) Formaldehyde (I) TCE

A close up of a logo

Description automatically generatedA close up of a logo

Description automatically generated

Moran’s *I* = 0.65 Moran’s *I* = 0.69 Moran’s *I* = 0.63

Figure 4: Local Clusters for Each Pollutant

A close up of a map

Description automatically generated

Moran’s *I* = 0.66

Figure 5: Local Clusters for Total Lung Cancer Incidence Rates

Table 2: Descriptive statistics for cancer incidence rates by race, gender, age and SESa index

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cancer Rate per 1,000 | Mean | Median | S.D.b | Range |
| Total | 10.74 | 9.69 | 6.11 | 0.00 – 75.00 |
| Race |  |  |  |  |
| White | 11.01 | 9.58 | 10.20 | 0.00 – 125.00 |
| African American | 16.37 | 0.00 | 38.77 | 0.00 – 333.33 |
| Gender |  |  |  |  |
| Male | 11.70 | 10.59 | 7.24 | 0.00 – 68.49 |
| Female | 10.09 | 9.09 | 8.51 | 0.00 – 125.00 |
| Race and Gender |  |  |  |  |
| White Male | 11.64 | 9.80 | 12.46 | 0.00 – 142.86 |
| White Female | 10.41 | 8.63 | 13.62 | 0.00 – 200.00 |
| African American Male | 18.09 | 0.00 | 63.36 | 0.00 – 1000.00 |
| African American Female | 18.30 | 0.00 | 75.29 | 0.00 – 1000.00 |
| Age |  |  |  |  |
| 40 – 49 Years | 0.94 | 0.00 | 1.90 | 0.00 – 17.24 |
| 50 – 59 Years | 7.25 | 5.23 | 8.85 | 0.00 – 111.11 |
| 60 – 69 Years | 17.45 | 14.41 | 14.69 | 0.00 – 142.68 |
| 70 – 79 Years | 19.56 | 17.27 | 15.31 | 0.00 – 172.41 |
| 80 and Over Years | 28.08 | 21.85 | 34.21 | 0.00 – 545.45 |
| Table 2 continued  SES |  |  |  |  |
| % African American | 16.19 | 3.50 | 26.37 | 0.00 – 98.00 |
| % Family Below Poverty | 12.98 | 6.10 | 50.16 | 0.00 – 100.00 |
| % Individuals Below Poverty | 13.81 | 9.10 | 13.51 | 0.00 – 100.00 |
| % At Least High School | 84.75 | 86.20 | 9.08 | 27.90 – 100.00 |

a Socioeconomic Status

b Standard Deviation

The aggregate lung cancer rates were strongly associated with coke over emissions (*p* = < 0.001) and TCE (*p* = 0.002) among all pollutants (Table 3).

Table 3: Association of lung cancer incidence rate adjusted with SES.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Coefficient (*β*) | Standard Error | *p*-value |
| 1,3- Butadiene | -21.336 | 45.741 | 0.641 |
| Acetaldehyde | 10.640 | 8.321 | 0.202 |
| Benzene | -3.824 | 2.563 | 0.136 |
| Chloroform | -10.709 | 7.969 | 0.180 |
| Coke Oven Emission | 20.914 | 5.811 | <0.001 |
| Diesel PM | 0.098 | 0.411 | 0.813 |
| Ethylbenzene | 14.193 | 9.291 | 0.127 |
| Formaldehyde | -7.003 | 6.133 | 0.254 |
| TCEa | 4.220 | 4.667 | 0.002 |

a Trichloroethylene

After log transformation (to achieve normality), linear regression revealed the relationship between lung cancer rates and exposure of coke oven emissions, and trichloroethylene by race, gender and age groups. TCE exposure was significantly associated with total lung cancer incidence rates (Table 4). SES, including the percentage of families below the poverty level and percentage of the population with at least high school education, were adjusted for in the analysis. Lung cancer in women seemed to be more strongly associated with coke oven emissions and TCE exposure than men (Table 4).

After adjusting for coke oven emissions and TCE, we found a statistically significant, positive relationship with percent families below poverty (*p*-value = <0.001) and negative relationship with educational attainment (*p*-value = <0.001). Population with less than high school education is at higher risk for lung cancer (Table 5).

Table 4: Association of exposure of pollutants and lung cancer variables adjusted for SESa

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coke Oven Emissionsb | | | Trichloroethyleneb | | |
| Standardized Incidence Rates | Coefficient (*β*) | t | *p*-value | Coefficient (*β*) | t | *p*-value |
| Total | 0.105 | 1.889 | 0.060 | 0.259 | 2.972 | 0.003 |
| Race |  |  |  |  |  |  |
| White | 0.028 | 0.505 | 0.614 | 0.286 | 3.305 | 0.001 |
| African American | 0.067 | 0.849 | 0.397 | 0.023 | 0.195 | 0.846 |
| Gender |  |  |  |  |  |  |
| Male | 0.052 | 0.939 | 0.348 | 0.052 | 0.939 | 0.348 |
| Female | 0.103 | 1.760 | 0.079 | 0.259 | 2.884 | 0.004 |
| Race and Gender |  |  |  |  |  |  |
| White Male | 0.024 | 0.413 | 0.680 | 0.169 | 1.891 | 0.059 |
| White Female | 0.041 | 0.687 | 0.493 | 0.268 | 2.901 | 0.004 |
| African American Male | 0.080 | 0.936 | 0.351 | 0.015 | 0.111 | 0.912 |
| African American Female | 0.014 | 0.158 | 0.874 | -0.091 | -0.702 | 0.484 |
| Age |  |  |  |  |  |  |
| 40 – 49 Years | 0.126 | 1.310 | 0.193 | 0.285 | 2.012 | 0.047 |
| 50 – 59 Years | 0.120 | 2.077 | 0.039 | 0.005 | 0.054 | 0.957 |
| 60 – 69 Years | 0.094 | 1.586 | 0.114 | 0.210 | 2.325 | 0.021 |
| 70 – 79 Years | -0.188 | -3.179 | 0.002 | 0.086 | 0.961 | 0.337 |
| 80 and Over Years | -0.069 | -1.131 | 0.259 | 0.110 | 1.165 | 0.245 |

a Socioeconomic Status

b explanatory variables

Table 5: Association of lung cancer and SESa, adjusted with Coke Oven Emissions and TCEa.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Coefficient (*β*) | t | *p*-value |
| Below Poverty - Families | 0.183 | 8.234 | <0.001 |
| Educational Attainment – At Least High School Education | -0.219 | -3.644 | <0.001 |

a Socioeconomic Status

b Trichloroethylene

## Conclusion

Over the past decade, there has been increasing awareness of cancer disparities in the United States. The federal government has initiated several programs to reduce the disparities such as ﻿CDC’s Racial and Ethnic Approaches to Community Health (REACH) and ﻿National Center on Minority Health and Health Disparities to lead and coordinate scientific efforts to improve the health of minorities and medically underserved people (Ward et al., 2004). Despite incredible efforts, knowledge is still limited for cancer prevention in vulnerable communities. This study attempted to identify the disadvantaged areas that are impacted the most from air pollution. This study utilized available cancer incidence data and exposure levels in Allegheny County to describe risk factors by specific areas. Nine pollutants were examined against lung cancer rates; trichloroethylene and coke oven emission were found to be the most associated with lung cancer. As described in the introduction, both pollutants emit in the environment from automobile exhaust and coke processing facilities. Further epidemiological studies will need to be undertaken to direct link these pollutants to lung cancer. However, occupational cohort studies have provided substantial evidence to assert trichloroethylene and coke oven emission as carcinogens (Pope III et al., 2002; Redmond, 1983; Scott & Jinot, 2011). This study also attempted to visualize exploratory spatial data on maps. Informative maps can provide comprehensive information to policymakers, researchers, and public health professionals in reducing air pollution and cancer prevention efforts.

This study’s results are also consistent with the previous studies that predicated probability of lung cancer increases with age (Jemal et al., 2008; Siegel, Miller, & Jemal, 2019b). About 80% to 90% of lung cancer deaths are linked to the consumption of tobacco products and cigarette smoking (Centers for Disease Control and Prevention, 2018). This study did not include variables such as smoking, occupation or other behavioral factors such as diet that may play a contributing role in lung cancer prevalence. Thus, this study may not be sufficient to affirm coke oven emission and TCE as sole contributors to lung cancer. Major limitations of this study are, the exposure was determined based on EPA’s modeled estimation; therefore, we assign exposure at a group level (census tract) but not individually. Indoor exposure for certain pollutants may be higher for some populations, such as those who rely on woodfire heat, are exposed to secondhand smoking, etc., which was not accounted in this study. Also, the migration to immigration ratio was ignored to maintain simplicity in the study. Seven pollutants in this study which were not found associated with lung cancer may suggest that their ambient presence may not pose any significant cancer risk. However, it is worth reevaluating the animal studies to identify exposure magnitude of these agents to humans by better modeling.

Future research could be done to study other cancer incidences as well as cancer mortality. The geospatial analysis could be done for other cancers to examine the spatial relationship between exposure and impact on health in vulnerable communities.

# bibliography

Agency for Toxic Substances and Disease Registry (CDC). (1997). Toxicological Profile: Chloroform. Retrieved April 6, 2019, from https://www.atsdr.cdc.gov/toxprofiles/tp.asp?id=53&tid=16

Agency for Toxic Substances and Disease Registry (CDC). (2010). Toxicological Profile: Ethylbenzene. Retrieved April 7, 2019, from https://www.atsdr.cdc.gov/toxprofiles/tp.asp?id=383&tid=66

Agency for Toxic Substances and Disease Registry (CDC). (2016). *National Toxicology Program Trichloroethylene (TCE) Exposure*. Retrieved from http://www.atsdr.cdc.gov/substances/

Aglan, M. A., & Mansour, G. N. (2018). Hair straightening products and the risk of occupational formaldehyde exposure in hairstylists. *Drug and Chemical Toxicology*. https://doi.org/10.1080/01480545.2018.1508215

American Cancer Society. (2014). Formaldehyde. Retrieved April 5, 2019, from https://www.cancer.org/cancer/cancer-causes/formaldehyde.html

American Cancer Society. (2016). Benzene. Retrieved April 6, 2019, from https://www.cancer.org/cancer/cancer-causes/benzene.html

Anselin, L. (1995). Local indicators of spatial organization -LISA. In *Geographical Analysis* (pp. 1–25). Ohio State University Press. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1538-4632.1995.tb00338.x

Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis*, *38*(1), 5–22. https://doi.org/10.1111/j.0016-7363.2005.00671.x

Avogbe, P. H., Ayi-Fanou, L., Autrup, H., Loft, S., Fayomi, B., Sanni, A., … Møller, P. (2005). Ultrafine particulate matter and high-level benzene urban air pollution in relation to oxidative DNA damage. *Carcinogenesis*, *26*(3), 613–620. https://doi.org/10.1093/carcin/bgh353

Bahadar, H., Mostafalou, S., & Abdollahi, M. (2014). Current understandings and perspectives on non-cancer health effects of benzene: A global concern. *Toxicology and Applied Pharmacology*, *276*(2), 83–94. https://doi.org/10.1016/j.taap.2014.02.012

Boffetta, P., & Nyberg, F. (2003). Contribution of environmental factors to cancer risk. *British Medical Bulletin*, *68*(1), 71–94. https://doi.org/10.1093/bmp/ldg023

Centers for Disease Control and Prevention. (2018). What Are the Risk Factors for Lung Cancer? Retrieved April 12, 2019, from https://www.cdc.gov/cancer/lung/basic\_info/risk\_factors.htm

Cohen, A. J. (2000). Outdoor Air Pollution and Lung Cancer Exposures to Carcinogens in Outdoor Air. *Environmental Health*, *108*(September 1999), 743–750. Retrieved from https://ehp.niehs.nih.gov/cms/attachment/45fb25e7-36e6-4185-aa9f-1554220ae02a/ehp.00108s4743.pdf

Fabisiak, J., Jackson, E., Brink, L., & Presto, A. (2018). *Environmental Health Perspectives A Risk-based Model to Assess Environmental Justice and Coronary Heart Disease Burden from Traffic-related Air Pollutants*.

Helbich, M., Leitner, M., & Kapusta, N. D. (2012). *Geospatial examination of lithium in drinking water and suicide mortality*. *International Journal of Health Geographics* (Vol. 11). https://doi.org/10.1186/1476-072X-11-19

Hemminki, K., & Pershagen, G. (1994). Cancer risk of air pollution: epidemiological evidence. *Environmental Health Perspectives*, *102 Suppl*(Suppl 4), 187–192. https://doi.org/10.1289/ehp.94102s4187

Huang, Y., Zou, Z., & Deng, H. (2007). An analysis report about peripheral blood anemia induced by excessive formaldehyde in abiding place. *Jiangsu Environ Sci Technol 2007b*, *20*, 16–17.

Janssen, N. A. H., Vliet, P. H. N. Van, Harssema, H., & Brunekreef, B. (2001). Assessment of exposure to traffic relatedair pollution of children attending schools near motorways. *Atmospheric Environment*, *35*(2), 3875–3884.

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., & Thun, M. J. (2008). Cancer Statistics, 2008. *CA: A Cancer Journal for Clinicians*, *58*(2), 71–96. https://doi.org/10.3322/CA.2007.0010

Lancianese, A. (2018, March 28). New Report Finds Industrial Pollution Flowing Illegally into PA Rivers - The Allegheny Front. Retrieved from https://www.alleghenyfront.org/new-report-finds-industrial-pollution-flowing-illegally-into-pa-rivers/

Lewtas, J. (2007). Air pollution combustion emissions: Characterization of causative agents and mechanisms associated with cancer, reproductive, and cardiovascular effects. *Mutation Research - Reviews in Mutation Research*, *636*(1–3), 95–133. https://doi.org/10.1016/j.mrrev.2007.08.003

Linder, S. H., Marko, D., & Ken, S. (2008). Cumulative cancer risk from air pollution in houston: Disparities in risk burden and social disadvantage. *Environmental Science and Technology*, *42*(12), 4312–4322. https://doi.org/10.1021/es072042u

Martenies, S. E., Milando, C. W., Williams, G. O., Batterman, S. A., Martenies, S. E., Milando, C. W., … Batterman, S. A. (2017). Disease and Health Inequalities Attributable to Air Pollutant Exposure in Detroit, Michigan. *International Journal of Environmental Research and Public Health*, *14*(10), 1243. https://doi.org/10.3390/ijerph14101243

Martien, P., Ph, D., Lau, V., Tanrikulu, S., & Ph, D. (2014). Identifying Areas with Cumulative Impacts from Air Pollution in the San Francisco Bay Area, (March). Retrieved from http://www.baaqmd.gov/~/media/Files/Planning and Research/CARE Program/Documents/ImpactCommunities\_2\_Methodology.ashx

Miller, J. H. R. M. H. S. J. H. M. P. R. (1984). Carcinogenicity of 1,3-Butadiene in B6C3F1 mice after 60 weeks of inhalation exposure. *Science*, *227*(4), 548–549.

National Institutes of Health. (2018). Cancer Statistics. Retrieved March 23, 2019, from https://www.cancer.gov/about-cancer/understanding/statistics

National Toxicology Program (NIH). (2016). *Report on Carcinogens, Fourteenth Edition Coke-Oven Emissions*. Retrieved from http://ntp.niehs.nih.gov/go/roc

Odum, J., Foster, J. R., & Green, T. (1992). A mechanism for the development of Clara cell lesions in the mouse lung after exposure to trichloroethylene. *Chemico-Biological Interactions*, *83*(2), 135–153. https://doi.org/10.1016/0009-2797(92)90042-J

Office of Environmental Health Hazard Assessment. (2001). *Health Effects of Diesel Exhaust*. Retrieved from www.californialung.org

Parent, M. É., Rousseau, M. C., Boffetta, P., Cohen, A., & Siemiatycki, J. (2007). Exposure to diesel and gasoline engine emissions and the risk of lung cancer. *American Journal of Epidemiology*, *165*(1), 53–62. https://doi.org/10.1093/aje/kwj343

Pennsylvania Department of Health. (2015). The Burden of Cancer in Pennsylvania.

Pennsylvania Department of Health. (2019). PA Environmental Justice Areas. Retrieved April 7, 2019, from https://www.dep.pa.gov/PublicParticipation/OfficeofEnvironmentalJustice/Pages/PA-Environmental-Justice-Areas.aspx

Pope III, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (2002). Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. *JAMA*, *287*(9), 1132. https://doi.org/10.1001/jama.287.9.1132

Redmond, C. K. (1983). Cancer mortality among coke oven workers. *Environmental Health Perspectives*, *VOL. 52*, 67–73. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1569361/pdf/envhper00458-0072.pdf

Salariino, A. J., Willey, J. C., Lechner, J. F., Grafstrom, R. C., Laveck, M., & Harris5, C. C. (1985). *Effects of Formaldehyde, Acetaldehyde, Benzoyl Peroxide, and Hydrogen Peroxide on Cultured Normal Human Bronchial Epithelial Cells*. *CANCER RESEARCH* (Vol. 45). Retrieved from http://cancerres.aacrjournals.org/content/45/6/2522.full-text.pdf

Samet, J. M., Marbury, M. C., & Spengler, J. D. (1987). State of Art: Indoor Air Pollution. *The American Review of Respiratory Disease*, (136), 1486–1508.

Sathiakumar, N., Brill, I., & Delzell, E. (2009). 1,3-butadiene, styrene and lung cancer among synthetic rubber industry workers. *Journal of Occupational and Environmental Medicine*, *51*(11), 1326–1332. https://doi.org/10.1097/JOM.0b013e3181c3c663

Scott, C. S., & Jinot, J. (2011). Trichloroethylene and cancer: Systematic and quantitative review of epidemiologic evidence for identifying hazards. *International Journal of Environmental Research and Public Health*, *8*(11), 4238–4272. https://doi.org/10.3390/ijerph8114238

Sharma, R. K., Stacy, S. L., Xue, T., Talbott, E. O., Brink, L. L., & Yuan, J.-M. (2019). *Neighborhood Disparities in Cancer: A Geospatial Analysis of Socioeconomic and Environmental Factors*.

Siegel, R. L., Miller, K. D., & Jemal, A. (2019a). Cancer statistics, 2019. *CA: A Cancer Journal for Clinicians*, *69*(1), 7–34. https://doi.org/10.3322/caac.21551

Siegel, R. L., Miller, K. D., & Jemal, A. (2019b). Cancer statistics, 2019. *CA: A Cancer Journal for Clinicians*, *69*(1), 7–34. https://doi.org/10.3322/caac.21551

Silverman, D. T. (2017). Diesel exhaust causes lung cancer: now what? *Occupational and Environmental Medicine*, *74*(4), 233–234. https://doi.org/10.1136/oemed-2016-104197

Singh, G. K., Williams, S. D., Siahpush, M., & Mulhollen, A. (2011). Socioeconomic, rural-urban, and racial inequalities in US cancer mortality: Part I-All cancers and lung cancer and part II-Colorectal, prostate, breast, and cervical cancers. *Journal of Cancer Epidemiology*, *2011*. https://doi.org/10.1155/2011/107497

Tang, X., Bai, Y., Duong, A., Smith, M. T., Li, L., & Zhang, L. (2009). Formaldehyde in China: Production, consumption, exposure levels, and health effects. *Environment International*, *35*(8), 1210–1224. https://doi.org/10.1016/j.envint.2009.06.002

Tomasetti, C., & Vogelstein, B. (2015). Variation in cancer risk among tissues can be explained by the number of stem cell divisions HHS Public Access. *Science*, *347*(6217), 78–81. https://doi.org/10.1126/science.1260825

Tornqvist, M., & Ehrenberg, L. (1994). On cancer risk estimation of urban air pollution. In *Environmental Health Perspectives* (Vol. 102, pp. 173–182). Retrieved from https://ehp.niehs.nih.gov/doi/pdf/10.1289/ehp.102-1566917

United States Environmental Protection Agency. (2000). *Chloroform*. Retrieved from https://www.epa.gov/sites/production/files/2016-09/documents/chloroform.pdf

United States Environmental Protection Agency. (2009). *1,3-Butadiene*. Retrieved from https://www.epa.gov/sites/production/files/2016-08/documents/13-butadiene.pdf

United States Environmental Protection Agency. (2016a). *Acetaldehyde*. Retrieved from https://www.epa.gov/sites/production/files/2016-09/documents/acetaldehyde.pdf

United States Environmental Protection Agency. (2016b). *Ethylbenzene*. Retrieved from https://www.epa.gov/sites/production/files/2016-09/documents/ethylbenzene.pdf

United States Environmental Protection Agency. (2017). Risk Management for Trichloroethylene (TCE). Retrieved from https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/risk-management-trichloroethylene-tce

United States Environmental Protection Agency. (2018). National Air Toxics Assessment. Retrieved April 8, 2019, from https://www.epa.gov/national-air-toxics-assessment

United States Environmental Protection Agency. (2019a). Environmental Justice. Retrieved from https://www.epa.gov/environmentaljustice

United States Environmental Protection Agency. (2019b). Toxics Release Inventory (TRI) Program. Retrieved April 7, 2019, from https://www.epa.gov/toxics-release-inventory-tri-program/learn-about-toxics-release-inventory

US Census Bureau. (2019). Geography Program. Retrieved April 7, 2019, from https://www.census.gov/programs-surveys/geography.html

Uttara, B., Singh, A. V, Zamboni, P., & Mahajan, R. T. (2009). *Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options*. *Current Neuropharmacology* (Vol. 7). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2724665/pdf/CN-7-65.pdf

Valavanidis, A., Vlachogianni, T., Fiotakis, K., & Loridas, S. (2013). Pulmonary oxidative stress, inflammation and cancer: Respirable particulate matter, fibrous dusts and ozone as major causes of lung carcinogenesis through reactive oxygen species mechanisms. *International Journal of Environmental Research and Public Health*, *10*(9), 3886–3907. https://doi.org/10.3390/ijerph10093886

Ward, E., Jemal, A., Cokkinides, V., Singh, G. K., Cardinez, C., Ghafoor, A., & Thun, M. (2004). Cancer disparities by race/ethnicity and socioeconomic status. *CA: A Cancer Journal for Clinicians*, *54*(2), 78–93. https://doi.org/10.3322/canjclin.54.2.78

Western Pennsylvania Regional Data Center. (2016). Allegheny County Environmental Justice Areas - Datasets. Retrieved April 7, 2019, from https://data.wprdc.org/dataset/environmental-justice-census-tracts

Western Pennsylvania Regional Data Center. (2019). The Region’s Open Data at Your Fingertips. Retrieved April 7, 2019, from http://www.wprdc.org/

White, M. C., Holman, D. M., Boehm, J. E., Peipins, L. A., Grossman, M., & Jane Henley, S. (2014). Age and cancer risk: A potentially modifiable relationship. *American Journal of Preventive Medicine*, *46*(3 SUPPL. 1). https://doi.org/10.1016/j.amepre.2013.10.029

Wilhelm, M., & Ritz, B. (2005). Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA. *Environmental Health Perspectives*, *113*(9), 1212–1221. https://doi.org/10.1289/ehp.7751

Yang, H. H., Lee, W. J., Chen, S. J., & Lai, S. O. (1998). PAH emission from various industrial stacks. *Journal of Hazardous Materials*, *60*(2), 159–174. https://doi.org/10.1016/S0304-3894(98)00089-2

1. Selected cancers are female breast cancer, uterine cervix, colon/rectum, uterine corpus, lung/bronchus, melanoma of the skin, non-Hodgkin lymphoma, prostate and urinary bladder. [↑](#footnote-ref-1)
2. The "Census Tract" is an area roughly equivalent to a neighborhood established by the Bureau of Census for analyzing populations. They generally encompass a population between 2,500 to 8,000 people. They are smaller than neighborhoods and bigger than blocks, and my not follow the city boundaries. Number of census tracts may vary over period of time. There were total 416 census tracts in 2000 and 402 in 2010 (US Census Bureau, 2019). [↑](#footnote-ref-2)
3. Common boundaries can be shared in Rook (only common edges) and Queen (both common edges and corners) contiguities. For precise estimation of data, we used Queen’s ﻿as a spatial contiguity. [↑](#footnote-ref-3)