Journal of Medicine and Public Health

ISSN: 2766-8355

Research Article

Analysis of NIH Statistical Cancer Maps and EPA Pollution Data Case Studies of Lung Cancer Clusters near Ports

Adwik Rahematpura[1], James Geller[1] and Soon Ae Chun[2]*

[1]New Jersey Institute of Technology, New Jersy, USA

[2]City University of New York College of Staten Island, Staten Island, New York, USA

Abstract

This analysis is based on the investigation of an NIH Statistical Cancer Maps dataset, and EPA pollution data. Using the cancer maps dataset, we discovered lung and bronchial cancer clusters of three or more counties, located near large ports. In the New Jersey map, we observed such a cluster. The problem we addressed was finding the possible reason for cancer clusters near ports and we hypothesized air pollution could lead to the clusters. Using EPA pollution data, we graphically demonstrated air quality measures and their recording locations. Furthermore, through visual analytics, we observed pollutant level spikes in areas near ports. Using wind data, we showed that in three NJ counties, shipping is highly correlated with increased cancer incidence. We supported this with three similar cases from other states. Public health officials should revisit the consequences of residential developments in proximity to ports and allowed time of attendance at such locations.

Keywords: Cancer; NIH; Lung cancer; Public health; CDC

Introduction

In this paper, we are raising the question why lung cancer clusters have been observed in areas near large ports. We analyzed an interactive map web page by the NIH and CDC, which displays cancer incidence projected onto maps [1], which we call "NIH Cancer Maps". Each map displays cancer incidence per 100,000 for county or state and is created based on multiple adjustable parameters. These parameters include type of cancer, age, demographic information, etc. As per the type of cancer, we chose to focus on lung and bronchial cancer. This was based on our initial hypothesis that the cancer clusters centralized near ports were due to air pollution. The International Agency for Research on Cancer (IARC), a part of the World Health Organization, "classified outdoor air pollution as a cancer-causing agent (carcinogen)" and "concluded that particulate matter causes lung cancer" [2,3].

The second dataset we used, US Pollution Data: Pollution in the U.S. since 2000 is composed of historical pollution data scraped from the EPA from 2000-2016 [4]. It includes data on four of the five main pollutants that the EPA uses to determine the Air Quality Index (AQI): Carbon Monoxide (CO), Nitrogen Dioxide (NO_2), Ground Level Ozone (O_3) and Sulfur Dioxide (SO_2). Data on particulate matter, the fifth major pollutant, is not in this dataset. IARC specifically mentions particulate matter as a lung cancer carcinogen, however, the other four main pollutants, included in the EPA dataset, are not classified as carcinogens as apparently no human trials have ever been performed.

For each gas in the above set, there are five items recorded: units (parts per million, parts per billion, etc.), mean (of all data points in a day), max value, hour of max value, and air quality index. The data was collected from a number of different sites around the United States of America. The dataset contains 1.75 million records and includes 29 measurement attributes: the 20 mentioned earlier (five items each for four gasses), as well as address, city, county, and state of the measuring station. We performed visual analytics to create multiple graphs, which will be shown and explained.

The third dataset we used was sourced from wind.willyweather.com and displays NOAA weather data [5]. The interactive website allowed us to enter specific cities and get current and historical wind averages. The questions we investigated were whether cancer clusters in NJ counties near ports could be attributed to air pollution under consideration of wind direction and speed, and if those reasons could be replicated in other states.

Background

We will now elaborate on the specifics of each pollutant gas being studied: carbon monoxide, nitrogen dioxide, ground level ozone and sulfur dioxide. Carbon monoxide comes from a variety of sources, including furnaces, leaking chimneys, generators, gasoline powered equipment, fuel vehicle exhaust etc [6]. According to the EPA, "Acute effects are due to the formation of carboxyhemoglobin in the blood, which inhibits oxygen intake. At moderate concentrations, angina, impaired vision, and reduced brain function may result. The National Ambient Air Quality Standard (NAAQS) sets two types of allowances: primary and secondary allowances. The primary allowance is defined as being for the general public and sensitive groups such as those with respiratory illnesses. The secondary allowance is for "welfare," protecting animals, buildings and crops [7]. Only a primary allowance is defined for carbon monoxide at 35 ppm over a one hour period.

Citation: Rahematpura A, Geller J, Chun SA. Analysis of NIH Statistical Cancer Maps and EPA Pollution Data Case Studies of Lung Cancer Clusters near Ports. J Med Public Health. 2023;4(8):1093.

Publisher Name: Medtext Publications LLC

Manuscript compiled: Nov 23rd, 2023

***Corresponding author:** Soon Ae Chun, City University of New York College of Staten Island, Staten Island, New York, USA

Nitrogen dioxide is, "...in the air from the burning of fuel. NO_2 forms from emissions from cars, trucks and buses, power plants and off-road equipment [8]. It is particularly harmful, because in small doses it can, "aggravate respiratory disease" and lead to long term damage. The primary and secondary allowances for NO_2 are the same at 53 ppb. Ground level ozone (O_3) should not be confused with the ozone layer found in the stratosphere [9]. Ground level ozone is a highly toxic gas that forms when, "pollutants emitted by cars, power plants, industrial boilers, refineries, chemical plants, and other sources chemically react in the presence of sunlight" [10]. Ozone is similar to nitrogen dioxide in that it also severely exacerbates pulmonary illness. A unique aspect about ozone is that since it is formed from the reactions of other pollutants and sunlight, it is the result of the other three gasses in this analysis. Hence, more of the other three gases likely results in higher levels of ground level ozone. According to the NAAQS table, the primary and secondary allowances for ozone are 0.07 ppm. Sulfur dioxide is an extremely harmful gas, which comes from the "burning of fossil fuels by power plants and other industrial facilities...natural sources such as volcanoes; and locomotives, ships and other vehicles and heavy equipment that burn fuel with high sulfur content" [11]. The primary allowance for SO_2 is 75 ppb and the secondary allowance is 0.5 ppm (500 ppb).m while air pollution is linked to lung cancer, it is also linked to liver/bile, breast and pancreatic cancer. It also impacts, "DNA repair function [alters] the body's immune response, or inflammation that triggers angiogenesis, the growth of new blood vessels that allows tumors to spread" [12].

Materials and Methods

Our methodology was based on using quantitative and visual analytics. It varied, based on the dataset. One dataset was sourced from the NIH website in the form of maps. The other dataset was from kaggle.com as a .CSV file. It allows the user to enter a number of parameters to create the map. We began with New Jersey. The parameters entered were as follows, in order: New Jersey, By County, Cancer Incidence, Lung and Bronchus, Incidence (All Stages), All Races, Both Sexes, All Ages, Latest Five Years.

The visual analytics were performed by importing geographic and air quality data into the Tableau software. Using the maps created with Tableau, we plotted the geo locations of air quality stations. These showed the changes in gas levels over time at different locations, covering all 50 states of the US. We are not showing Hawaii in the maps. We also created maps displaying the average levels of each pollutant for the entire US between 2000 and 2016. We focused on and analyzed five geographic cancer clusters and different sources of air pollution in their proximity. This included five ports (Wilmington DE, Savannah GA, Seattle/Tacoma WA, Los Angeles CA, and Virginia, VA). We also considered industrial sources of air pollution, such as DuPont Chemicals in Salem County, NJ, and volcano eruptions in Washington State. For industrial sources we perused news reports, legal documents and general information websites [13,14]. However, the latter two sources did not provide useful additional data points and were eventually dropped from the analysis.

We charted meteorological data to show how pollutants could have been blown to large cancer cluster areas. For this, we used wind data from willyweather.com, a website which displays NOAA weather data5. For information on shipping and fuel, we used data from Notteboom and Cariou15. This included information such as fuel usage at different speeds of ships. If a cancer cluster was close to a port, but covered a much larger geographic area, we used wind data in an effort to explain the extension direction of the cluster.

Results

Analyzing the map (Figure 1), a cluster of three counties stands out: Gloucester County, Salem County, and Cape May County. Those three counties have the highest rates of lung and bronchial cancer per 100,000 in the state of New Jersey. Gloucester County had 75.5 per 100,000, Salem County had 82 per 100,000, and Cape May County had 81.4 per 100, 0001. We observed that the cluster was in close proximity to the Port of Wilmington (Delaware). Our initial hypothesis was that this cancer cluster was due to high air pollution levels from the port.

The Delaware City Refinery, one of the largest inputs of oil on the east coast and is a major port with many incoming container ships. The most common fuel for large container ships is called bunker fuel [16]. Currently, pending further regulations, the maximum level of sulfur content in bunker fuel is 3.5% however its average is 2.7% according to ExxonMobil [17]. Fuel consumption in ships, measured in tons, is based on the TEU (Twenty-foot Equivalent). This represents one container. According to data from the Institute of Transport and Maritime Management, a container ship of 8000-9000 TEU at 21 knots consumes 146.2 tons of bunker fuel in a day [15]. Notably, 21 knots is a relatively slow speed used due to entering the Delaware River and due to the container ship's proximity to the port. At 21 knots, a ship at full capacity would burn 3.94 tons of sulfur into sulfur dioxide per day based on a fuel sulfur content of 2.7%. The Delaware River defines the western border of New Jersey, and it can be inferred that the high lung cancer incidence trend could not continue further north up the river as after Camden County (directly north of Gloucester County), there are no more ports for the ships to go to. The Delaware River also becomes too narrow for shipping. Similar analyses apply for Cape May County. It lies just at the mouth of the Delaware River where ships are still going at their normal speed and potentially burning more sulfur than near Salem County.

To investigate whether there was a high level of sulfur in the air surrounding the Port of Wilmington, we used the EPA Pollution dataset. Simultaneously, we also investigated the air quality at four other major ports in the US: The Port of Savannah, Seattle/Tacoma, Los Angeles, and Virginia. We created a map showing the locations of recording sites (Figure 2). It shows that there is a recording site in Wilmington, Delaware where the port is, as well as in Seattle, Virginia, and Los Angeles near their respective ports. There is a recording site in Camden, NJ, which is in close proximity to the Port of Wilmington as well. The map did not show a measuring station in Georgia, however, the air quality maps still displayed data, and therefore we still investigated the Port of Savannah. As the focus was on sulfur dioxide, we created a map displaying its level over a period from 2000 to 2016 (Figure 3).

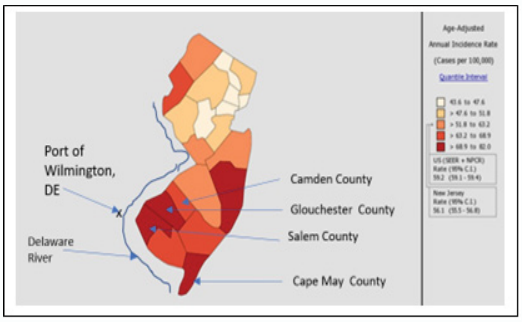

Figure 1: Lung and bronchial cancer incidence in New Jersey by county.

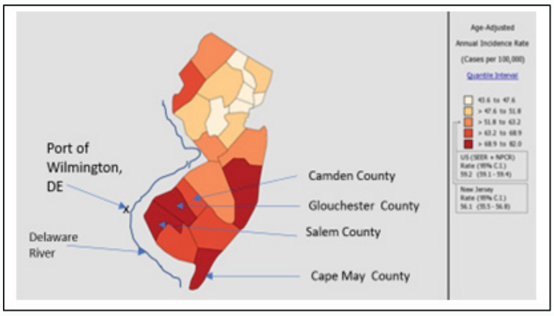

Figure 2: Locations of air quality recording sites.

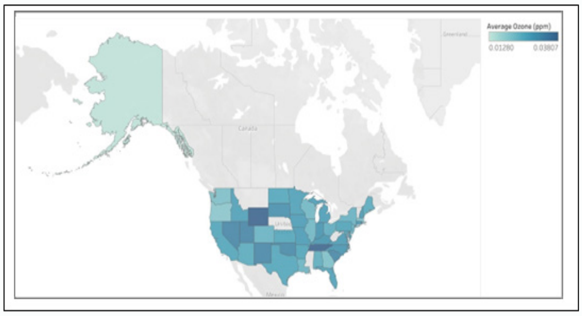

Figure 4: Average sulfur dioxide levels from 2000 to 2016.

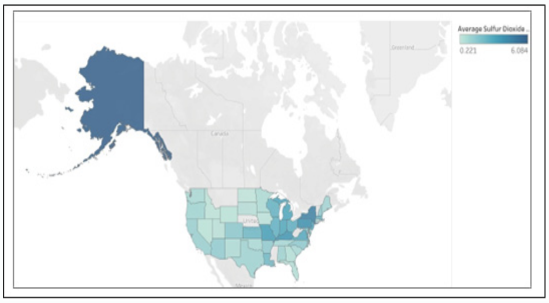

Figure 3: Average sulfur dioxide levels from 2000 to 2016.

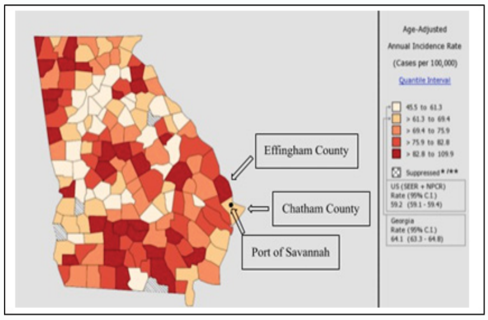

Figure 5: Average sulfur dioxide levels from 2000 to 2016.

Figure 4 shows a map displaying data for ground level ozone from 2000 to 2016 as, it is a byproduct of sulfur dioxide in contact with sunlight. The maps show that for Delaware, the levels were of average size, however, for neighboring New Jersey they were fairly high, as were the ozone levels. The levels for Seattle and Georgia were also in the mid-range. California had extremely low levels of both Sulfur Dioxide and Ozone, and Virginia had very high levels of both. We observed that the clusters were generally grouped in areas that were in close proximity to the ports. Using wind data, we investigated whether the pollutants could have been blown in from the port/recording site, to the cluster area. Beginning with the Port of Savannah, we created a map of lung and bronchial cancer incidence using the same methodology as in the map for New Jersey. The port of Savannah is the fourth largest in the United States. We edited the resulting map to have an arrow showing where the port of Savannah is located (Figure 5). We observed that the counties due west of the port indicate high levels of lung and bronchial cancer.

Contrastingly, the county in which the Port of Savannah is situated, Chatham County has low rates. We again hypothesized that the reason for this could be wind. According to meteorological data, Chatham county generally has a constant west/northwest facing wind speed of about 5-8 mph [18]. This is a light breeze that is likely forceful enough that the sulfur dioxide and other gasses emitted by cargo ships will be blown away from Chatham county and the neighboring counties. One neighboring county is Effingham county about 30 miles away to the northwest, which has relatively high levels of lung and bronchial cancer at 84.7 per 100,000 people. For reference, the lowest incidence rate in a county in Georgia is 45.5 per 100,000 people, and the highest is 97.1 per 100,000.The Port of Seattle/Tacoma is the fifth largest in the United States. As seen in the map (Figure 6), the area of the port is indicated. Just south of that area is a set of counties (Pierce County, Lewis County, Thurston County) that have high lung and bronchial cancer incidence. Seattle has strong southward winds that have averaged about 6.7 mph since 2010 [19]. The case of Los Angeles is an outlier and will be discussed in a later section.

The Port of Wilmington, a port that is large but not as large as the ones discussed above, did follow a similar pattern. Due east of the port across the Delaware River is an area of New Jersey with high levels of lung and bronchial cancer. Data shows that Wilmington has historically short hard bursts of wind westward, but the steady wind is about 7-12 mph due East [20]. This would hypothetically push air pollutants to the counties in New Jersey previously mentioned. Finally, for the Port of Virginia. The evidence is not as clear, and a definitive wind report was not found, but counties near the port did experience abnormally high levels of lung and bronchial cancer (Figure 7).

High levels of sulfur dioxide and ozone were also observed. A potential reason could be due to a large naval base near the port with personnel being there for extended periods of deployment. When we observed potential cancer clusters near Los Angeles, the counties surrounding the Port of Los Angeles had among the lowest lung and bronchial cancer incidences in the country (Figure 8). This appeared to be an outlier. We checked the historical wind direction data which was generally due south west, away from the other counties [21]. However, eventually, we found that the Port of Los Angeles has had a plan called the Clean Air Action Plan since 2006, which greatly reduced pollution from the port [22]. The Port of Savannah only adopted a similar plan from the EPA in 2012. While these plans are directed at the ports themselves and not the ships, it is possible that this is why the Port of Los Angeles has very low lung and bronchial cancer incidence.

In the EPA Pollution Dataset, there is one distinct limitation. In the EPA dataset, there were many air quality recording sites in some states, while other states had only one, or sometimes zero sites. Furthermore, some states were not surveyed each year between 2000 and 2016. This results in more accurate measurements for states with

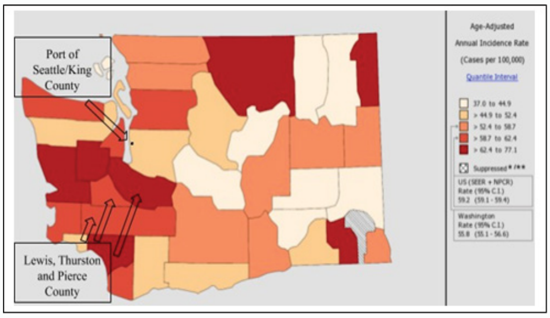

Figure 6: Average sulfur dioxide levels from 2000 to 2016.

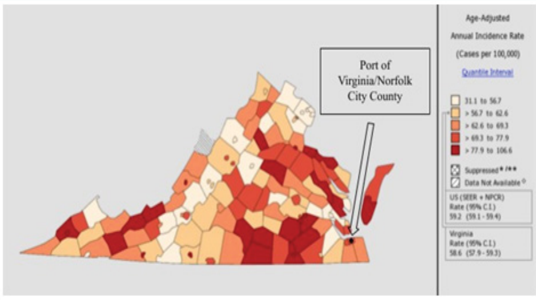

Figure 7: Average sulfur dioxide levels from 2000 to 2016.

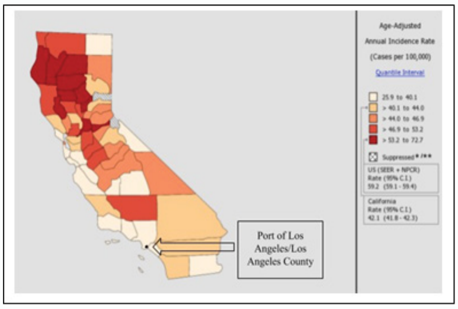

Figure 8: Average sulfur dioxide levels from 2000 to 2016.

more measurement sites and years recorded, and less accurate data for those with few sites and shorter recording times. For states with only one site, the single site is not an accurate representation of the whole state.

Related work

To the best of our knowledge, there are only two papers/analyses that are similar in purpose, methodology and/or conclusion to this paper. The first related paper is from Wine brake et al [23]. This paper focuses on two main ideas: the rules and regulations for sulfur content in fuel, and their respective health impacts as modeled by various statistical models [23]. The authors begin by explaining the different scenarios they used which were a "no control 2.7% sulfur content fuel," a 0.5% and 0.1% sulfur content limit within 200 nautical miles of a coastline and a 0.5% global limit. The four scenarios were tested using the global climate model: ECHAM5/MESSy1-MADE. All of the results in their paper were based on the results from that statistical model, which showed that the number of deaths for the three scenarios. The research also considered policies such as the International Maritime Organization and SECA, the Sulfur Emission Control Area where the sulfur content in fuel is extremely limited. All results in their study are based on a model that showed that mortality will likely decrease due to a reduction in sulfur content in fuel. The study also only considered one major pollutant which was PM 2.5 which is particulate matter with a maximum diameter of 2.5 micrometers [23]. Our study focuses on the four other major pollutants.

A second related analysis was performed by Mitchel, of the California Environmental Protection Agency [24]. The main points were regarding potential health problems specifically labeling the lung as the major organ affected. Mitchel analyzed different air pollutants as well. We know that Mitchel used news reports and medical databases, as well as statistical analysis. Mitchel's study identifies Diesel as the primary concern, whereas we focus on the general concept of bunker fuel. Throughout the report, references are made to lung disease and conditions such as asthma or simple inflammation, however, no reference is made to cancer [24].

Conclusions

Using visual analytics, we have shown that the integration of different data sources, such as lung cancer cluster data, major ports, air pollution dataset and weather data, can reveal some insights that lung cancer clusters may be related to their proximity to major ports. It is not possible to prove causation based on such co-occurrence data. However, in the interest of population health, this phenomenon should receive more attention in the future.

Wind data, including average direction and speed, should also be taken into account, because areas where pollutants are being blown to would enlarge the actual geographic area at risk. Allowed time of attendance at ports, or locations near them, should also be considered. Prolonged exposure to pollutants from ports appears to be harmful to health and has to be balanced with desirable economic activity. Concerning future work, public health officials should (re) consider laws and guidelines regarding the proximity of new housing developments to ports. Ports provide large numbers of jobs at many different skill levels. Furthermore, where there is a port, there is usually a beach, and housing close to beaches is also considered desirable to many people. Thus, there would be a natural inclination to develop housing near a port, both to cut down on commuting time and to enable spare time activities at the ocean, like fishing, boating, swimming, etc. However, in light of the health hazards that were observed, new housing in such areas might not be advisable.

Acknowledgements

This work is partially sponsored by NSF 1747728, and National Research Research Foundation of Korea (NRF- 2017S1A3A2066084).

References

1. Interactive maps. State cancer profiles. 2020.
2. Simon S. World health organization: outdoor air pollution causes cancer. American Cancer Society; 2013.
3. Editorial Staff. The connection between lung cancer and outdoor air pollution. American Lung Association; 2018.
4. US Environmental Protection Agency. U.S. pollution data. Environmental Protection Agency; 2020.
5. Willy Weather. USA weather forecast. Willy Weather; 2020.
6. US EPA, Office of Air and Radiation. Carbon monoxide's impact on indoor air quality. US EPA; 2018.
7. US EPA, Office of Air and Radiation. NAAQS table. US EPA; 2018.

8. US EPA, Office of Air and Radiation. Basic information about NO2. US EPA; 2016.

9. US EPA. The ozone problem, ground-level ozone. US EPA; 2018.

10. US EPA, Office of Air and Radiation. Ground-level ozone basics. US EPA; 2015.

11. US EPA, Office of Air and Radiation. Sulfur dioxide basics. US EPA; 2016.

12. Exposure to particulate air pollutants associated with numerous cancers. Science Daily; 2016.

13. DuPont offices and global locations. Dupont; 2010.

14. Mordock J. N.J. town files $1 billion lawsuit against DuPont. The News Journal; 2016.

15. Notteboom T, Cariou P. Fuel surcharge practices of container shipping lines: is it about cost recovery or revenue-making. In: International Association of Maritime Economics. Copenhagen; 2009.p.24-26.

16. Penner D. What is bunker fuel? Vancouver Sun; 2015.

17. What does IMO's 0.50% sulphur cap decision will affect bunker supply chain. ExxonMobil; 2019.

18. Willy Weather. Chatham county wind forecast, GA. Willy Weather; 2019.

19. Willy Weather. Seattle wind forecast, WA. Willy Weather; 2020.

20. Willy Weather. Wilmington wind forecast, DE. Willy Weather; 2020.

21. Willy Weather. Los Angeles wind forecast, CA. Willy Weather; 2020.

22. The Port of Los Angeles. San pedro bay ports clean air action plan. 2010.

23. Winebrake JJ, Corbett JJ, Green EH, Lauer A, Eyring V. Mitigating the health impacts of pollution from oceangoing shipping: an assessment of low-sulfur fuel mandates. Environ Sci Technol. 2009;43(13):4776-82.

24. Mitchell D. Health effects of shipping related air pollutants. In: EPA Region 9 Conference on Marine Vessels and Air Quality. California; 2001.