Benzene Exposure, Leukemia, and Non-Hodgkin’s Lymphoma in a Spatial Context

Rebecca Shyu1, Timothy Haltinholt2, and Iris Zachary1

1Electrical Engineering and Computer Science Department, Informatics Institute, 2Health Management and Informatics Department
University of Missouri – Columbia, MO 65211, USA

Abstract

This research looks at various environmental benzene sources: tobacco smoke, plastic industries, petroleum industries, and benzene-related chemicals released in industries and its known association with leukemia and Non-Hodgkin’s lymphoma in geospatial context. It uses geospatial analytics to understand relationships between the environment and leukemia & non-Hodgkin lymphoma. Our hypothesis is that the greater the number of high environmental benzene sources, the higher incidence rate for each cancer. We found two factors, specifically plastic industries and benzene-related chemicals with statistically significant relationships with leukemia and non-Hodgkin lymphoma incidence rates.

Introduction/Aims

Benzene, a classified carcinogen by IARC, NTP, and EPA, is known to cause leukemia [1]. It is widely used in industries and is among the top 20 chemicals in production volume in the United States due to its role as a starting chemical for plastics and dyes. Cigarette smoke is responsible for over 50% of benzene exposure. Many studies have been done on industry workers with frequent, direct exposure to benzene. However, there are not many studies focusing on benzene exposure through the environment for the common population. Non-Hodgkin lymphoma has been suspected to be linked to benzene, but it is not widely established [2]. This study looks to find data-driven evidences of a geospatial relationship between benzene exposure from the environment and incidence rates on the county level for leukemia and Non-Hodgkin lymphoma.

Methods & Results

Benzene and Non-Hodgkin Lymphoma:

- Source: Cancer Control Planet (SEER and NPCR-SS) dataset
- No data from Kansas, Minnesota, and Nevada due to state legislation
- Plot incidence rates by quantile
- 2011-2015 Average Leukemia Incidence Rates (Figure 1)
- Non-Hodgkin Lymphoma Incidence Rates (Figure 2)

Leukemia & Non-Hodgkin Lymphoma

- Amount of chemical disposed separated by quantile (Figure 5)
- Result (Leukemia): Mean of Q1 set = 12.8571, Mean of Q4 set = 12.4021
- T = 4.1661, df = 1392.2, P-value = 3.2894e-05

Smoking Rates separated by quantile (Figure 3)
- Result (Leukemia): Mean of Q1 set = 9.773224, Mean of Q4 set = 12.91635
- T = 1.8424, df = 1421, P-value = 0.06563

Result (NHL):
- Mean of Q4 set = 12.91635
- T = 1100.4, df = 99.306, P-value = 2.886e-26

Result (Leukemia):
- Mean of Q4 set = 11.02
- T = 8.9295, df = 126.15, P-value = 4.293e-15

Conclusion and Discussions

From our results, we found two environmental benzene sources that had statistically significant p-values. Plastic’s and Benzene-Related Chemicals’ Q1 and Q4 data sets showed noticeable differences in the means of incidence rates for both leukemia and Non-Hodgkin lymphoma. Others, such as smoking and the petroleum industry did not support our hypothesis. This work demonstrates the potential use of environmental factors in geospatial analytics for cancer epidemiology. Future work includes expanding to multiple cancer types, other environmental sources, and social & economic factors.

Acknowledgements

The authors would like to thank the Honors College at the University of Missouri for funding RS through the undergraduate Discovery Fellow program. TH was funded by the National Institutes of Health under award number ST32GM121450. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References


Contact Info: Rebecca Shyu, RYSF001@missouri.edu