

Association of Social Determinants of Health with Respiratory Disease: A County-by-County Analysis for Ohio

Rohan M. Desai

Hawken School, 12465 County Line Rd, Chesterland, OH 44026, USA; desairohan2005@gmail.com

Mentors: Raoul R. Wadhwa, Jacob G. Scott

ABSTRACT: A recent research focus on social determinants of health has revealed their importance in the healthcare setting. Even prior to the COVID-19 pandemic, about 20% of United States zip codes experienced economic distress, marked by high poverty rates and unemployment. These areas are particularly vulnerable to adverse clinical outcomes. Here, we studied clinical outcomes of respiratory diseases in the state of Ohio, specifically COVID-19 and lung cancer, and how they were affected by social determinants of health, specifically median county-level income. Three publicly available datasets were merged, and associations between pairs of variables were evaluated using Spearman's rank correlation coefficient and Student's t-test. Our analysis found that clinical outcomes of both respiratory diseases were significantly correlated with median county-level income ($p < 0.001$). This implies that although variation between disease outcomes exists across Ohio, social determinants of health play a common, underlying role. As a result, public health intervention across Ohio is necessary to support health equity. Similar studies should be undertaken for other states and territories within the United States to determine the need for public health intervention across the country.

KEYWORDS: Biomedical and Health Sciences; Other; Cancer; COVID-19; Social Determinants of Health.

■ Introduction

Despite significant improvements in diagnosis and management of sub-acute and chronic diseases, there continue to be challenges of unexplained residual risk accounting for disparate outcomes. In recent times, community-level socioeconomic distress has been identified as a potential driver of outcomes following many common diseases, including common respiratory illnesses like chronic obstructive pulmonary disease and lung cancer.¹⁻³ Unfortunately, a fifth of United States zip codes currently demonstrate social and economic distress, including high poverty rates, joblessness, and economic recession.⁴ Measures of community distress such as low education levels, low median household income, high rates of poverty, and air quality are important determinants of its residents' overall health and well-being.^{5,6} Consequently, patients with respiratory diseases living in socioeconomically distressed neighborhoods have a higher risk of adverse clinical outcomes.

Since 2019, the world has been dealing with the novel coronavirus disease 2019 (COVID-19), which predominantly manifests with acute respiratory symptoms. Interestingly, the aforementioned social determinants of health (SDOH) might play a key role in determining outcomes, even considering a relatively novel and acute disease like COVID-19. The state of Ohio represents a bell-weather state in the USA with a blend of socioeconomic strata, political ideologies, and urban versus rural areas. Whereas most published research uses national-level data that potentially obfuscate associations due to the presence of unrelated variables, such as varying culture and diet, using state-level data permits a more reasonable assumption of homogeneity with regard to cultural variables. We hypothesized that, in the state of Ohio, there was a significant

association between household income, an important surrogate for SDOH, and acute (COVID-19 infection) or sub-acute (lung cancer) respiratory illnesses. Furthermore, we predicted that higher household income would be associated with lower COVID-19 fatality and lower incidence of lung cancer.

■ Methods

Median county-level household income was used as a proxy for social determinants of health. The severity of COVID-19 was measured by case fatality, defined as the percentage of confirmed cases of COVID-19 with a fatal outcome. Lung cancer incidence was defined as the number of newly diagnosed cases per 100,000 members of the population. The proxy variables were chosen based on reasonable relation to the outcome we intended to measure⁷⁻¹⁰ and based on the availability of datasets. Association between these variables was measured with Spearman's rank correlation coefficient, which measures the association between two quantitative variables based solely on order instead of magnitude. The statistical significance of the pairwise correlations was evaluated with Student's t-test, which permits statistical testing of a predetermined hypothesis. The false-positive threshold was set to 5%, or 0.05; a two-tailed test was used in accordance with the principle of conservatism. Although relying on clinically relevant data, the study only used publicly available and aggregated datasets. As such, it is exempt from IRB approval under category 4: secondary research of publicly available data without identifiable information.

The current study utilized 3 publicly available datasets: (1) county-level median income, which was collected from the United States Census Bureau; (2) county-level lung cancer incidence, which is publicly available on the National Cancer Institute website; and (3) county-level COVID-19 case fatality

ratios, downloaded from Johns Hopkins University's publicly available repository.¹¹⁻¹⁴ All datasets contained complete information for Ohio's counties ($n = 88$) and were merged by county name with no name conflicts or missing values. Data manipulation, visualization, and the statistical inference were conducted in the R programming language. The R code used to conduct the statistical analysis for this manuscript is available via a publicly accessible GitHub code repository at the following link: <https://www.github.com/desairohan2005/covid-lung-sdoh>.

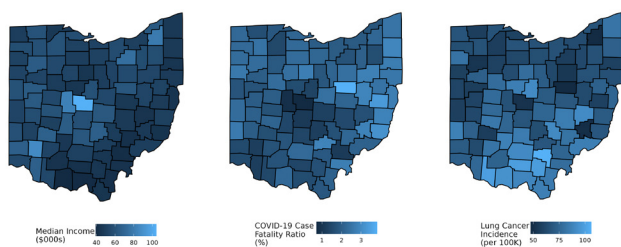


Figure 1: Distribution of median income, COVID-19 case fatality ratio, and lung cancer incidence in counties of Ohio. Higher (lighter color) median county-level income (left) visually corresponds to lower (darker color) county-level COVID-19 case fatality ratio (center) and lower (darker color) county-level lung cancer incidence (right).

■ Results and Discussion

Results:

Median county-level income across Ohio was \$52,723 (IQR: 47,274-59,167). As a result, the median county-level lung cancer incidence was 68.4 (IQR: 62.0-80.6) per 100,000 population, and the median COVID-19 case fatality ratio was 1.98% (IQR: 1.65-2.36). The distribution of the above parameters on a county-by-county basis is shown in Figure 1. A relationship between median income (Figure 1, left) and both of the SDOH variables (Figure 1, center and right) is visually apparent. Specifically, counties with a higher median income (lighter color) also have decreased COVID-19 case fatality (darker color) and decreased lung cancer incidence (darker color). Data for 7 counties with the highest and lowest median income each are shown in Table 1. There also appears to be a clear pattern within the numerical data in this table. That is, the 7 counties with the highest median incomes tend to have lower values for COVID-19 case fatality and lung cancer incidence in comparison with the 7 counties that have the lowest median incomes in the state of Ohio.

The patterns deduced from the figure and table were confirmed quantitatively. The two SDOH variables, lung cancer incidence and COVID-19 case fatality were not significantly correlated ($r = -0.20, p = 0.1$). In contrast, median income was significantly correlated with both lung cancer incidence ($r = -0.42, p < 0.001$) and COVID-19 case fatality ($r = -0.49, p < 0.001$). This implies that 24% of the variation in lung cancer incidence and 18% of the variation in COVID-19 case fatality ratio in counties across Ohio was explained by median county income. Scatterplots of all three pairwise associations are included as supplemental figures at <https://www.github.com/desairohan2005/covid-lung-sdoh>.

Significance of Results:

We have demonstrated that as the median income of an Ohio county increases, the likelihood of dying due to COVID-19 and being diagnosed with lung cancer decreases, suggesting that SDOH plays a crucial role in clinical outcomes of 2 distinct respiratory diseases, one acute and another subacute/chronic. Interestingly, COVID-19 case fatality and lung cancer incidence were not significantly associated, thereby reducing the likelihood of confounding. Previous studies have explained how management, in addition to incidence, of lung cancer is affected by SDOH.^{2,3} We validated these studies at the state-level and demonstrated the need for public health intervention.

We also encourage incorporating SDOH consideration into clinical care. This can be done qualitatively and quantitatively within healthcare settings. For instance, qualitative interventions would include emphasizing health counseling by primary care physicians. Although primary care physicians serve large patient populations, they can focus their efforts based on the specific SDOH vulnerabilities present in the area they serve. On the other hand, quantitative interventions could include the incorporation of SDOH variables into clinical decision-support tools. In particular, risk calculators have gained traction within clinical care. Yet, despite the clear effect of SDOH on clinical outcomes, no risk calculators currently use SDOH variables in predicting outcomes. Our study clearly indicates that this is a limitation that future risk calculators should overcome.

Table 1: Counties in Ohio with the highest and lowest median incomes (seven of each). Corresponding values for lung cancer incidence

Rank	County	Median Income (\$000s)	Lung cancer incidence (per 100K)	Risk of lung cancer relative to median	COVID-19 case fatality (%)	Risk of fatality relative to median
1	Delaware	104.3	51.4	0.75	0.71	0.36
2	Warren	87.1	62.9	0.92	1.24	0.63
3	Union	82.8	66.2	0.97	0.85	0.43
4	Geauga	80.0	48.0	0.70	2.20	1.11
5	Medina	74.2	56.9	0.83	1.73	0.87
6	Greene	67.1	56.5	0.83	1.62	0.82
7	Clermont	67.0	75.9	1.11	1.24	0.63
82	Highland	43.3	73.5	1.07	1.74	0.88
83	Pike	42.8	99.0	1.45	1.42	0.72
84	Morgan	42.3	80.5	1.18	2.20	1.11
85	Gallia	42.1	83.7	1.22	2.06	1.04
86	Athens	40.9	74.4	1.09	1.13	0.57
87	Scioto	39.7	88.7	1.30	1.59	0.80
88	Adams	39.1	93.6	1.37	2.25	1.14

Strengths and Limitations:

Data reliability is a strength of this study. All the data sources were established organizations that make their data publicly available. This not only increases data reliability by opening it to public critique but also increases the reproducibility of the study. Thus, researchers and reviewers of this study can themselves replicate the results to validate quantitative outputs. Notably, as time passes, researchers can also replicate this study in the future to determine whether the effects of SDOH over time are changing. This feature will be particularly useful after implementing public health interventions to determine whether they have had a valuable impact on healthcare outcomes.

A limitation of this study is the county-level aggregation. Many SDOH studies have been conducted at the national level, which limits their application to individual states in a large, heterogeneous country like the United States. Although our study limits this issue by focusing on county-level data within a single state, counties themselves still possess a measure of heterogeneity that is lost by aggregation. However, this aggregation is a necessary limitation, as having less aggregation would risk confidentiality. Additionally, publicly available data is generally limited to county-level or zip code-level aggregation, both of which contain group sizes of similar magnitude.

Future Directions:

We are particularly interested in two extensions of our study. First, this cross-sectional analysis of Ohio should be repeated individually for other states and territories within the United States. Given the level of funding and autonomy provided to state governments, individual studies are warranted as public health interventions could be more effective if customized to target the SDOH vulnerabilities in each state. Second, this cross-sectional analysis could be extended to a longitudinal one. At the same time, we evaluate the association between SDOH and clinical outcomes at a specific point in time; these associations represent complex relationships that will likely evolve over time due to a multitude of reasons, such as legislative changes or public health programs. Thus, comparing our results to the associations a few years in the future will provide valuable insight into understanding whether health inequity is growing or decreasing over time within the state of Ohio.

Conclusion

We have shown that social determinants of health are significantly associated with respiratory disease-related health outcomes within the state of Ohio. For example, patients diagnosed with COVID-19 are more likely to die because of it when they reside in a lower-income county. Additionally, people living in lower-income counties are more likely to be diagnosed with lung cancer. Although the numerical results might vary across the United States, it is likely that our results generalize across the United States. Therefore, public health interventions are necessary to manage the potential detriment to health caused by social determinants and address healthcare inequity.

Acknowledgments

The authors thank the Hawken School, Cleveland Clinic, and Case Western Reserve University for their support. This was a no-cost, unfunded study.

References

1. Pakhale, S.; Tariq, S.; Huynh, N.; Jama, S.; Kaur, T.; Charon, C.; Florence, K.; Nur, F.; Bustamante-Bawagan, M. E.; Bignell, T.; Boyd, R.; Haddad, J.; Kendzerska, T.; Alvarez, G. Prevalence and Burden of Obstructive Lung Disease in the Urban Poor Population of Ottawa, Canada: A Community-Based Mixed-Method, Observational Study. *BMC Public Health* **2021**, *21*, 183.
2. Takemura, T.; Kataoka, Y.; Okazaki, K.; Sakurai, A.; Imakita, T.; Ikegaki, S.; Matsumoto, H.; Saito, E.; Takata, H.; Kaku, S.; Wada, N.; Shinomiya, M.; Otoshi, T.; Shimada, M.; Nikaide, J.; Iki, R.; Hirano, K.; Hirai, T.; Endo, K.; Naganuma, T. Influence of Social Determinants of Health on Patients with Advanced Lung Cancer: A Prospective Cohort Study. *BMJ Open* **2018**, *8*, e023152.
3. Namburi, N.; Timsina, L.; Ninad, N.; Ceppa, D.; Birdas, T. The Impact of Social Determinants of Health on Management of Stage I Non-small Cell Lung Cancer. *Am. J. Surg.* **2021**, *223* (6), 1063-1066.
4. Economic Innovation Group. The Distressed Communities Index Report. Economic Innovation Group, 2016. <https://eig.org/wp-content/uploads/2020/10/EIG-2020-DCI-Report.pdf> (accessed 2022-07-01).
5. Thornton, R. L.; Glover, C. M.; Cene, C. W.; Glik, D. C.; Henderson, J. A.; Williams, D. R. Evaluating Strategies for Reducing Health Disparities by Addressing the Social Determinants of Health. *Health Aff.* **2016**, *35*, 1416-1423.
6. Forchuk, C.; Dickins, K.; Corring, D. J. Social Determinants of Health: Housing and Income. *Healthc. Q.* **2016**, *18*, 27-31.
7. Valencia, C. I.; Gachupin, F. C.; Molina, Y.; Batai, K. Interrogating patterns of cancer disparities by expanding the social determinants of health framework to include biological pathways of social experiences. *Int. J. Environ. Res. Public Health* **2022**, *19* (4), 2455-2465.
8. Jou, K.; Sterline, M. R.; Ramos, R.; Antoine, F.; Nanus, D. M.; Phillips, E. Eliciting the social determinants of cancer prevention and control in the catchment of an urban cancer center. *Ethn Dis.* **2021**, *31* (1), 23-30.
9. Green, H.; Fernandez, R.; MacPhail, C. The social determinants of health and health outcomes among adults during the COVID-19 pandemic: A systematic review. *Public Health Nurs.* **2021**, *38* (6), 942-952.
10. Rodriguez, J.; Quintana, Y. Understanding the social determinants of health and genetic factors contributing to the differences observed in COVID-19 incidence and mortality between underrepresented and other communities. *J. Natl. Med. Assoc.* **2022**, *114* (4), 430-439.
11. United States Census Bureau. <https://www.census.gov> (accessed 2022-04-20).
12. National Cancer Institute. <https://www.cancer.gov> (accessed 2022-04-22).
13. Center for Systems Science and Engineering, Johns Hopkins University. COVID-19 Data Repository. GitHub, 2022. <https://www.github.com/CSSEGISandData/COVID-19> (accessed 2022-04-15).
14. Dong, E.; Hongru, D.; Gardner, L. An Interactive Web-based Dashboard to Track COVID-19 in Real Time. *Lancet Infect. Dis.* **2020**, *20* (5), 533-534.

Authors

Rohan M. Desai is a high school student in Ohio interested in battling healthcare inequity in the United States. He will attend Vanderbilt University, where he will major in Medicine, Health, and Society.

Raoul R. Wadhwa, MD, recently completed medical school in Ohio. He is interested in primary care medicine.