Population-Level Disease Prevalence Rates Correlate With COVID-19 Mortality

George L. Morris, III, MD, MPH

ABSTRACT

Introduction: Initial reports identified preexisting conditions associated with COVID-19 mortality risk. The Centers for Disease Control and Prevention (CDC) 500 Cities Project provides prevalence rate estimates at the census tract level for these conditions. The frequency of these individual condition prevalence rates may associate with the census tracts with greater risk of COVID-19 deaths.

Objective/Research Question: Can the census tract-level outcome of Milwaukee County COVID-19 death rates correlate with the census tract-level COVID-19 individual mortality risk condition prevalence rates?

Methods: This study used the 296 Milwaukee County, Wisconsin census tracts' COVID-19 death rates per 100,000 lives to perform a linear regression with individual COVID-19 mortality risk condition prevalence rates, obtained from the CDC's 500 Cities Project, and a multiple regression with 7 condition prevalence rates. The Milwaukee County Medical Examiner provided census tract identified deaths from COVID-19 from March 2020 through May 2020. Crude death rates for these 3 months per 100,000 population were analyzed in a multiple linear regression versus prevalence rates for these conditions in each census tract.

Results: There were 295 assessable COVID-19-related deaths in Milwaukee County in early 2020. The model of crude death rates showed statistical significance with the condition prevalence rates in Milwaukee County. A regression analysis of each condition's prevalence rate showed no association with crude death rates.

Conclusions: This study supports a correlation between high COVID-19 mortality rate census tracts and prevalence rate estimates of conditions associated with high individual COVID-19 mortality rates. The study is limited by the small COVID-19 death sample and the use of a single location. The ability to focus COVID-19 health promotion may save future lives if mitigation strategies are applied extensively in these neighborhoods.

• • •

Author Affiliation: School of Public Health, Imperial College London (Morris).

Corresponding Author: George L. Morris III, MD, MPH; email george.morris@ascension.org; ORCID ID 0000-0002-8930-1437

INTRODUCTION

The COVID-19 pandemic challenged public health and medical providers.[1] A fatal COVID-19 clinical course has been associated with preexisting factors, including age group, race, sex, obesity, hypertension, diabetes mellitus, cardiovascular disease, chronic pulmonary disease, and chronic kidney disease.[2] Milwaukee County neighborhoods in Wisconsin contain a wide prevalence rate range of these individual factors.[3] The disparities in COVID-19 death rates in urban settings, like Milwaukee County, are clustered among minorities and in neighborhoods known for higher prevalence rates of these conditions.[4]

Health interventions, such as social distancing, wearing face masks, and good hygiene, have been projected to significantly reduce COVID-19 transmission rates.[5] In addition, vaccine awareness and uptake has predicted a decline in COVID-19 cases in the United States.[6]

Critical health issue interventions, like health promotion for the COVID-19 pandemic, are provided in a decentralized fashion at local health departments in the US. Local data within a health department's responsible area are valuable to create tailored localized health communication.[7] The specificity of health data at a US city level (hundreds of thousands of people) or ZIP code level (tens of thousands of people) can have considerable variability, such as seen in New York City, New York.[8]

Table 1. Mean Prevalence and Range for the Seven Relevant Disease Factors Among Tracts

Factors Among Tracts	Mean Prevalence %	Range %
Chronic asthma	11.10	7.9–16.30
Coronary heart disease	6.05	2.00–12.30
Chronic obstructive pulmonary disease	6.67	2.30–12.80
Chronic renal disease	3.17	1.20–6.80
Diabetes mellitus	11.13	3.00–25.00
Hypertension	31.04	11.80–51.00
Obesity	37.24	3.00–53.70

Table 2. Model and Factor Association Values

Model	95% CI	*P* value
All conditions	-88.20 to 134.63	0.001
Chronic asthma	-12.13 to 11.72	0.973
Coronary heart disease	-40.61 to 20.98	0.531
Chronic obstructive pulmonary disease	-20.29 to 20.02	0.989
Chronic renal disease	-14.21 to 109.62	0.131
Diabetes mellitus	-13.97 to 10.30	0.766
Hypertension	-2.84 to 1.83	0.671
Obesity	-4.50 to 2.03	0.456

The US census is conducted every 10 years, and information is gathered in smaller geographic regions, known as census tracts. Data at census tract levels involve single neighborhoods (hundreds to several thousand people) and are more similar as a group than larger, ZIP code-level areas.[8] This similarity among smaller groups at the census tract level versus the larger ZIP code-level populations allows specific health issue interventions, like COVID-19 mitigation and vaccination education, for the relevant population need.[9]

Health data using small area estimates are needed but uncommon.[8] The CDC's 500 Cities Project, supported by the Robert Wood Johnson Foundation, provides prevalence rate estimates of 27 health-related issues at the census tract level based on responses to the Behavioral Risk Factor Surveillance Survey (BRFSS).[9] These issues include prevention, behaviors, and disease measures and are presented as prevalence by census tract across the United States. Health outcome measures in the 500 Cities estimates include cancer, stroke, arthritis, mental health, teeth loss, chronic asthma, coronary heart disease, diabetes, hypertension, chronic renal disease, and obesity.[9] These prevalence estimates have been confirmed with samples of national and local prevalence rates.[10,11] The measures were developed to assist with assessments of and planning for health interventions at the more granular, neighborhood level.[12] These measures have highlighted the significant disparities at the neighborhood level and the need to address disparities at that granular layer.[7] The 500 Cities prevalence data have not been confirmed as a source to predict outcomes, such as COVID-19 death rates.

The value of census tracts can be in their focused size. Kong and Zhang have documented the greater homogeneity in smaller area analysis.[7] The authors document ZIP codes as having more heterogenicity than neighborhood-level health data. For an example of the size issue, the city of Milwaukee ZIP code–53206–contains multiple census tracts and over 22,000 residents, while the largest census tract in Milwaukee County contains just over 6,000. The data provided by the 500 Cities Project have many applications. Bu et al and Perlman encourage seeing this data as a planning tool.[13,14] Copello et al have applied these prevalence data to planning for disease management.[15] This planning involves different hospitals and clinics and is aided by small census tract samples, and its use has been verified by other data sources, such as medical services billing.[7]

This study asks the research question: Can the census tract-level outcomes of Milwaukee County COVID-19 death rates correlate with the census tract-level COVID-19 individual mortality risk condition prevalence rates? The aim is to show that the neighborhoods with the highest COVID-19 death rates are associated with higher disease prevalence rates at the census tract level in Milwaukee County, Wisconsin.

METHODS

Data Sources

Health Prevalence Data: The CDC's 500 Cities Project "provided city- and census tract-level small area estimates for chronic disease risk factors, health outcomes, and clinical preventive services use for the largest 500 cities in the United States. These small area estimates allowed cities and local health departments to better understand the burden and geographic distribution of health-related variables in their jurisdictions and assisted them in planning public health interventions."[11]

Health outcomes prevalence rates include chronic asthma (CASM), coronary heart disease (CHD), chronic obstructive pulmonary disease (COPD), diabetes mellitus (DM), hypertension (HTN), chronic renal disease (CRD), and obesity (OBS). The dataset contains the relevant population of Milwaukee County residents at the census tract level (adults 17 and older) and all of the relevant disease prevalence rates associated with greater COVID-19 mortality.

COVID-19 Mortality Data: Deaths occurring in Milwaukee County from COVID-19 were tracked and provided on public request from the Milwaukee County Coroner's Office. These data contain age, sex, and census tract locations for the months March 2020 through May 2020. The coroner's office collects these data for reporting to the Wisconsin Department of Health for the National Bureau of Vital Statistics.

Analysis

Data from the 500 Cities dataset and COVID-19 mortality data were combined into a single Excel comma-separated values file and imported for analysis into the open-source statistical package version 3.6.3 with base R statistical packages (R Foundation for Statistical Computing, Vienna, Austria).

The R code is: ModelCDRandCM<-lm(CDR~CASM+CHD+COPD+CRD+DM+HTN+OBS), where

1) Crude death rate (CDR) is defined as assessable deaths per 100,000 lives in each of 296 census tracts and
2) the 7 CPR (CASM, CHD, COPD, CRD, DM, HTN, and OBS) are the mean prevalence rates in percentages for each illness.

The assumptions of linearity, variance, independence, and normality were established by reviewing the data prior to their inclusion in the analysis. The data were reviewed in a scatter plot and in a review of the residuals for variance and distribution. The alpha value was set at 95%.

RESULTS

The Milwaukee County Medical Examiner reported 368 deaths from COVID-19 from March 2020 through May 2020. Analysis was performed on 295 deaths of Milwaukee County residents. The deaths excluded from the analysis included those not part of Milwaukee County Census Tracts (N=28) and those recorded as nursing home residents in the tract (N=45), as residents of these care facilities have unknown tract origins.

Table 1 demonstrates the mean 7 disease prevalence rates and the prevalence range across the 296 Milwaukee County census tracts. The CDR for the 296 Milwaukee County census tracts were 29.48/100,000 (range 0–224.92).

Table 2 shows the results of the model with COVID-19 crude death rates as the outcome and the disease prevalence rate percentages as multiple factors for the regression. The model with the diseases' prevalence rates showed statistical significance for the total deaths (95% CI, -88.20 to 134.63; $P < 0.001$).

The individual census tracts were not associated with death rates nor were they associated with individual disease prevalence rates ($P > 0.05$)

DISCUSSION

This study found that higher COVID-19 mortality rates were associated with the 500 Cities estimated prevalence of 7 COVID-19 mortality risk conditions at a census tract level in Milwaukee County. The model statistically significantly associated COVID-19 death rates with all these conditions together, despite the early and small COVID-19 death sample. Individual disease prevalence rates separately did not correlate with higher COVID-19 mortality. The specificity for the community COVID-19 deaths reinforces how the health data characteristics of the neighborhood can associate with the neighborhood health outcomes. This association could support using the data from the CDC 500 Cities Project to plan areas for higher priority health issue intervention.

The COVID-19 pandemic highlights health promotion planning issues. First, the 500 Cities Project data were used to project the COVID-19 impact at the city level, ZIP code level, and for individuals. Du et al used the 500 Cities estimates to identify Texas city-level medical resource needs.[16] Do and Frank used ZIP code-level data to identify how communities of color are disproportionately affected by COVID-19 mortality.[17] Their conclusions about "neighborhoods" using the ZIP codes are based on a predominance racial percentage. They identified those ZIP codes with the predominant race of White, Black, Hispanic, and Other with populations over 40,000 to 60,000 persons in group.

The demographic refinement of census tract-level information may have presented even stronger associations for the study. Jin et al created a web-based calculator for COVID-19 individual risk using some of the 500 Cities prevalence estimates.[18] Bu et al asserts that data should be linked to critical outcomes and for planning and effective intervention.[13] This study links COVID-19 death rates to a valuable outcome in identifying high COVID-19 mortality risk neighborhoods.

Additionally, the health data collected by BRFSS provided great value for a COVID-19 pandemic that was not foreseen. This study shows how the 500 Cities prevalence estimates could and may have allowed strategic interventions at a neighborhood level. Strategic planning was important, as many unexpected burdens fell on local health departments to do more with the same resources. Many health departments, including Milwaukee County, used these risk-producing diseases' maps to identify at-risk populations. This study's result suggests that disease risk association assumption with COVID-19 outcomes may have been appropriate.

Study Limitations

This study's small sample size provides statistically significant findings but with wide confidence intervals. A larger sample is currently under investigation for March 2020 through May 2021. The sample is valid only in Milwaukee County, but the 500 Cities project reaches many cities where additional confirmation could be found. Milwaukee County's significant health disparities made the county a desirable choice for study, despite the smaller sample.[3]

Second, the 500 Cities data are an estimate of disease prevalence rates based on BRFSS surveys. The validity of the projections has been confirmed in 2 studies, but warnings about using projections have been made.[10,11,19] This study does use 2 independent data sources—the 500 Cities Project and Milwaukee County Medical Examiner's report—that address these warnings.

Race, age, and socioeconomic status (SES) factors have been associated with higher COVID-19 mortality rates.[2,17] The study did find that the individual disease prevalence rates were not predictive. Race, age, and SES may not independently be correlated but collectively – as with the preexisting conditions – add to the strength of the correlation. The model strength may be enhanced by including the census tract racial percentages, median ages, educational attainment, and income measures.

There are estimate limitations to the 500 Cities Project. Estimates are for adults 17 years and older and residents of urban areas. Childhood deaths did not appear in the first months of the pandemic, so no comparison would have been possible. Rural communities often are not collected in health data and have significant health access issues to make more health data collection valuable.

Finally, the disease prevalence percentages are a very simplistic approach to modelling a correlation. Demographic features, such as age, race, and poverty, were also part of individual risk factors for COVID-19 mortality. This report was originally intended to answer the question regarding whether conditions that posed increased mortality risk for individuals could predict population-level mortality events. More work and identified features likely will be necessary to find a well-fitted model.

CONCLUSIONS

Disease prevalence estimates that correlate death outcomes may help local health departments direct health promotion resources. Larger samples, inclusion of other parameters, and wider application to other communities would demonstrate a wider use for this approach. The COVID-19 pandemic illustrates that health data collection may have many unexpected health promotion planning benefits. Funding for health data survey expansion could improve public health resource use and health promotion effectiveness.

Funding/Support: None declared.

Financial Disclosures: None declared.

REFERENCES

1. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. *JAMA*. 2020;324(8):782-793. doi:10.1001/jama.2020.12839

2. Wortham JM, Lee JT, Althomsons S, et al. Characteristics of persons who died with COVID-19 - United States, February 12-May 18, 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(28):923-929. doi:10.15585/mmwr.mm6928e1

3. Greer DM, Baumgardner DJ, Bridgewater FD, et al. Milwaukee Health Report 2013: Health Disparities in Milwaukee by Socioeconomic Status. Center for Urban Population Health; 2014. Accessed August 25, 2021. https://institutionalrepository.aah.org/pop/1/

4. Dorn AV, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. *Lancet*. 2020;395(10232):1243-1244. doi:10.1016/S0140-6736(20)30893-X

5. Ferguson NM, Laydon D, Nedjati-Gilani G, et al. *Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand*. Imperial College London; 2020. Report 9. Accessed April 25, 2023. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf

6. Makhoul M, Ayoub HH, Chemaitelly H, et al. Epidemiological impact of SARS-CoV-2 vaccination: mathematical modeling analyses. *Vaccines (Basel).* 2020;8(4):668. doi:10.3390/vaccines8040668

7. Kong AY, Zhang X. The use of small area estimates in place-based health research. *Am J Public Health*. 2020;110(6):829-832. doi:10.2105/AJPH.2020.305611

8. Bi Q, He F, Konty K, Gould LH, Immerwahr S, Levanon Seligson A. ZIP Code-level estimates from a local health survey: added value and limitations. *J Urban Health.* 2020;97(4):561-567. doi:10.1007/s11524-020-00423-z

9. Division of Population Health, National Center for Chronic Disease Prevention and Health Promotion. 500 Cities Project: 2016 to 2019. Updated December 8, 2020. Accessed April 25, 2023. https://www.cdc.gov/places/about/500-cities-2016-2019/index.html

10. Zhang X, Holt JB, Lu H, et al. Multilevel regression and poststratification for small-area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. *Am J Epidemiol.* 2014;179(8):1025-1033. doi:10.1093/aje/kwu018

11. Klompas M, Cocoros NM, Menchaca JT, et al. State and local chronic disease surveillance using electronic health record systems. *Am J Public Health.* 2017;107(9):1406-1412. doi:10.2105/AJPH.2017.303874

12. Dorn AV, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. *Lancet.* 2020;395(10232):1243-1244. doi:10.1016/S0140-6736(20)30893-X

13. Bu DD, Liu SH, Liu B, Li Y. Achieving value in population health big data. *J Gen Intern Med.* 2020;35(11):3342-3345. doi:10.1007/s11606-020-05869-0

14. Perlman SE. Use and visualization of electronic health record data to advance public health. *Am J Public Health*. 2021;111(2):180-182. doi:10.2105/AJPH.2020.306073

15. Copello E, Mease A, Umapathy K, et al. Developing a GIS dashboard tool to inform non-profit hospitals of community health needs at the neighborhood level. *J Inf Syst Appl Res.* 2020;13(2):33-53.

16. Du Z, Wang X, Pasco R, et al. COVID-19 Healthcare Demand Projections: 22 Texas Cities. UT COVID-19 Consortium; 2020. Accessed August 24, 2021. https://sites.cns.utexas.edu/sites/default/files/cid/files/covid-19_analysis_for_22_texas_cities_033020.pdf

17. Do DP, Frank R. Unequal burdens: assessing the determinants of elevated COVID-19 case and death rates in New York City's racial/ethnic minority neighbourhoods. *J Epidemiol Community Health.* 2020;jech-2020-215280. doi:10.1136/jech-2020-215280

18. Jin J, Agarwala N, Kundu P, et al. Individual and community-level risk for COVID-19 mortality in the United States. *Nat Med.* 2021;27(2):264-269. doi:10.1038/s41591-020-01191-8

19. Ogburn EL, Rudolph KE, Morello-Frosch R, Khan A, Casey JA. A warning about using predicted values from regression models for epidemiologic inquiry. *Am J Epidemiol.* 2021;190(6):1142-1147. doi:10.1093/aje/kwaa282

advancing the art & science of medicine in the midwest

WMJ

WMJ (ISSN 1098-1861) is published through a collaboration between The Medical College of Wisconsin and The University of Wisconsin School of Medicine and Public Health. The mission of *WMJ* is to provide an opportunity to publish original research, case reports, review articles, and essays about current medical and public health issues.