Risk factors associated with mortality of COVID-19 in 3125 counties of the United States

Ting Tian[1#], Jingwen Zhang[1#], Liyuan Hu[1#], Yukang Jiang[1#], Congyuan Duan[1], Zhongfei Li[5], Xueqin Wang[2,3]*, Heping Zhang[4]*

Affiliations:

[1]School of Mathematics, Sun Yat-sen University, Guangzhou, Guangdong, China
135 Xingang Xi Road, Guangzhou, Guangdong, 510275, China

[2]School of Statistics, Capital University of Economics and Business, Beijing, China
121 Zhangjialukou, Huaxiang Fengtai District, Beijing, 100070, China

[3]School of Management, University of Science and Technology of China, Hefei, Anhui, China
96 Jinzhai Road, Hefei, Anhui, 230026, China

[4]School of Public Health, Yale University, New Haven, CT, USA
60 College Street, New Haven, CT, 06520-8034, USA

[5]School of Management, Sun Yat-sen University, Guangzhou, Guangdong, China
135 Xingang Xi Road, Guangzhou, Guangdong, 510275, China

[#]Dr. T. Tian, J. Zhang, L. Hu and Y. Jiang contributed equally to this article

*Dr. X. Wang and Dr. H. Zhang are corresponding authors

Abstract

Background

The number of cumulative confirmed cases of COVID-19 in the United States has risen sharply since March 2020. A county health ranking and roadmaps program has been established to identify factors associated with disparity in mobility and mortality of COVID-19 in all counties in the United States.

Methods

To find out the risk factors associated with county-level mortality of COVID-19 with various levels of prevalence, a negative binomial design was applied to the county-level mortality counts of COVID-19 as of August 27, 2020 in the United States. In this design, the infected counties were categorized into three levels of infections using clustering analysis based on time-varying cumulative confirmed cases from March 1 to August 27, 2020. COVID-19 patients were not analyzed individually but were aggregated at the county-level, where the county-level deaths of COVID-19 confirmed by the local health agencies.

Results

3125 infected counties were assigned into three classes corresponding to low, median, and high prevalence levels of infection. Several risk factors were significantly associated with the mortality counts of COVID-19, where higher level of air pollution (0.153, $P<0.001$) increased the mortality in the low prevalence counties and elder

individuals were more vulnerable in both the median and high prevalence counties . The segregation between non-Whites and Whites and higher Hispanic population had higher likelihood of risk of the deaths in all infected counties.

Conclusions

The mortality of COVID-19 depended on sex, race/ethnicity, and outdoor environment. The increasing awareness of the impact of these significant factors may lead to the reduction in the mortality of COVID-19.

Key words: adverse health factors; county-level confirmed and deaths; race/ethnicity; segregation index; physical environment

Introduction

COVID-19 is an infectious disease caused by a novel coronavirus with an estimated average incubation period of 5.1 days[1]. It spreads through person-to-person transmission, and has now infected 215 countries and regions with over 24 million total confirmed cases as of August 27, 2020[2]. The United States had 5,867,785 confirmed cases on August 27, 2020, the highest in the world, but there were only 69 confirmed cases on March 1, 2020[3].

The United States has been suffering from a severe epidemic, with COVID-19 related deaths occurring all over the country. For instance, New York City had the largest number of total deaths (23,674), accounting for the majority of deaths in the infected counties, while no one in King county, Texas was infected as of August 27, 2020[3]. Therefore, it is of great interest to find out the risk factors that influence the number of deaths of COVID-19. It is known that infectious diseases are affected by factors other than medical treatments[4, 5]. For example, influenza A is associated with obesity[6], and the spread of the 2003 SARS events depends on seasonal temperature changes[7].

The County Health Rankings and Roadmaps program was launched by both the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute[8]. This program has been providing annual sustainable source data including health outcomes, health behaviors, clinical care, social and economic factors, physical environment and demographics since 2010. We explored putative

risk factors that may affect the mortality of COVID-19 in different areas of the United States in order to increase awareness of the disparity and aid the development of risk reduction strategies.

Methods

Data sources

We collected the number of cumulative confirmed cases and deaths from March 1 to August 27, 2020, for counties in the United States from the New York Times[9]. The COVID-19 confirmed cases and deaths were identified by the laboratory RNA test and specific criteria for symptoms and exposures from health departments and U.S. Centers for Disease Control and Prevention (CDC). The county health rankings reports from year 2020 were compiled from the County Health Rankings and Roadmaps program official website[8]. There were 77 measures in each of 3142 counties, including the health outcome, health behaviors, clinical care, social and economic factors, physical environment, and demographics. We refer to the official website of the County Health Rankings and Roadmaps program[8] for detailed information.

Study areas

As of August 27, 2020, a total of 3,208 counties reported confirmed cases in the United States, leaving 3125 counties with both confirmed cases of COVID-19 and

county health ranking data recorded to be analyzed in this study. The total number of deaths as of August 27, 2020 was considered as the outcome of this study.

Assessment of covariates in health factors

We divided the putative risk factors[8] into 5 categories: health behaviors (e.g., access to exercise opportunities, insufficient sleep), clinical care (e.g. primary care physicians ratio), social and economic factors (e.g., racial segregation index), physical environment (e.g., transit problems and air quality), and demographics (age, sex, rural, and race/ethnicity). For example, there were previous studies which identified the air pollution may relate to high levels of COVID-19[10] and elder population had the high risk in the COVID-19[11]. Besides these identified risk factors, we were interested in the adverse health factors may link to the mortality of COVID-19. Table 1 presented descriptive definition, sources and literature of 12 risk factors. All deaths resulted from complications of COVID-19.

Statistical analysis

The trend of the cumulative confirmed cases varied greatly in counties of the United States. We used the partitioning around medoids (PAM) clustering algorithm[12, 13] to assign counties with similar trends into a homogenous class after standardizing the time series of cumulative confirmed cases from March 1 to August 27, 2020. Based on the clustering results, we used the Kruskal-Wallis test[14] to

detect whether there were significant differences in the distributions of 12 risk factors across different classes of counties. The 12 risk factors were used to build a negative binomial model[15, 16] for every class of the counties. The analysis was conducted in R version 3.6.1.

Validation analysis

We randomly divided counties (samples) into training (70% of the counties) and testing (30% of the counties) in each class. The model obtained from the training data was employed to predict the death counts of COVID-19 in the testing data, and the accuracy was assessed by the root mean square error (RMSE) of the mortality ratio (the number of deaths divided by the number of cumulative confirmed cases).

Results

Three classes of county-level infection in the United States

The clustering analysis grouped the 3,125 counties were assigned into 3 classes. There were 2,751 counties in the first class with the lowest overall cumulative confirmed cases. Its medoid was Halifax County in Virginia. There were 294 counties in the second class with a median level of overall cumulative confirmed cases. Its medoid was St. Clair County in Illinois. There were 80 counties in the third class with the highest overall cumulative confirmed cases. Its medoid was Marion County in Indiana. Here, the PAM algorithm selected the county with most representative data

as the medoid in a class[12, 13]. The geographical distribution of the counties by class was shown in Figure 1, where the size of a circle indicated the cumulative confirmed cases on August 27, 2020. The distribution of deaths on August 27, 2020, which clearly differed among the three classes, was also presented in Figure 1. Note that the east, south, and west coasts were the most severely hit areas by COVID-19. Most counties in the high prevalence class were from Massachusetts, New York, New Jersey, Florida, Texas and California [9].

Distributions of 12 selected risk factors in the three classes of counties

Figure 2 showed the distributions of the 12 selected risk factors by the class of counties. The distributions were significant different ($P<0.001$) for all 12 risk factors. For example, the average population in the low prevalence class was 38,444, which was 10% and 3% of the average populations in the median and high prevalence classes, respectively. The average proportion of rural residents in the low prevalence class was 64.47%, versus 2.72% in the high prevalence class. The segregation index of non-Whites versus Whites was the largest in the high prevalence class, but the smallest in the low prevalence class.

Factors influencing mortality of COVID-19 in the three classes

There were three common factors, namely, residential segregation between non-Whites and Whites, resident population, and the percentage of Hispanic population,

which had statistically significant (P<0.05) effects on mortality in all classes. The negative binomial model was used to understand the within-class effects of residential segregation between non-Whites and Whites and the percentage of Hispanic population on mortality of COVID-19 as shown in Figure 3. Note that the higher values of both residential segregation between non-Whites and Whites and the percentage of Hispanic population the higher mortality of COVID-19. In the high prevalence class, an increase in both the residential segregation between non-Whites and Whites and the percentage of Hispanic population resulted in more deaths than other two classes of counties.

Table 2 presented the significant factors specific to each class based on the training data. Specifically, in the low prevalence class, nine variables were significantly associated with the mortality of COVID-19. Higher values in the average daily density of $PM_{2.5}$ (0.153, P<0.001), the percentage of workforce driving alone to work (0.039, P<0.001), the percentage of workforce that had more than 30 minutes commute driving alone (0.015, P<0.001), the percentage of adults who reported less than average 7 hours sleeping (0.073, P<0.001), resident population (P<0.001), the percentage of Hispanic (0.020, P <0.001) and female population (0.054, P<0.001) and segregation index (0.015, P<0.001) , significantly increased the number of deaths, while more people living in rural areas (-0.014, P<0.001) decreased the number of deaths of COVID-19.

In the median prevalence class, eight variables were significantly associated with the deaths of COVID-19. Higher values in the percentage of workforce driving alone to work (0.016, P=0.008), the percentage of workforce that had more than 30 minutes commute driving alone (0.013, P<0.001), resident population (P<0.001), the percentage of population aged over 65 (0.049, P<0.001), the percentage of Hispanic (0.020, P<0.001) and female population (0.127, P=0.001) and segregation index (0.025, P<0.001)-led to an increase in deaths as opposite to a decrease in deaths of COVD-19 for more people living in rural areas (-0.011, P=0.007) .

In the high prevalence class, four variables were significantly associated with mortality. Higher values in resident population (P<0.001), , the percentage of population aged over 65 (0.114, P<0.001), the percentage of Hispanic population (0.014, P=0.009) and segregation index (0.019, P=0.005), caused more deaths.

For each class of counties, the model obtained from the training data was employed to predict the deaths of COVID-19 on August 27, 2020 using the testing data. The corresponding RMSE values for the mortality ratio were 0.056%, 0.041%, and 0.088%, respectively, in the low, median, and high prevalence classes.

Discussion

Using the time trends of the cumulative confirmed cases in 3,125 counties in the United States, we categorized those counties into three levels of infection. The low prevalence class counted for 88% of the 3,125 counties. Their resident population was

remarkably smaller than the other two classes of counties. But the resident population size increased the mortality of COVID-19 regardless of the level of COVID-19 prevalence. A higher population density may increase more contacts in social distancing[17, 18], leading to a higher risk in mortality of COVID-19. On the contrary, a higher percentage of residents living in rural areas in both the low and median prevalence classes of counties may reduce the mortality. Disparities in race and ethnicity were found in the infected populations. For example, Blacks were reported to be prone to COVID-19[19, 20], and living settings of racial/ethnic minorities were founded to be more crowded, making social distancing difficulty [21]. In this study, we found that Hispanics were more vulnerable. Further investigation is warranted to study the racial disparity in the mortality of COVID-19. However, the segregation index between non-Whites and Whites revealed the racial disparity in health, leading to differences in health status not only at the individual level but also at the community level[22]. A higher values in the segregation index indicated the poor health status, which may increase the mortality of COVID-19[17]. This health inequality increased the mortality rates of COVID-19 in all classes of counties.

For the low and the median prevalence class of counties, more workforce driving alone to work and commuting long-distance may increase the levels of anxiety[23], leading to the high mortality in COVID-19. A higher percentage of long-distance commuting workforce was also linked to a high level of anxiety for commuters[23]. And substantial time spent by long-distance commuters could inhibit their healthy

behaviors[24]. The stress and less healthy behaviors may increase individual's vulnerability to COVID-19[27-29]. Also, long-distance commute may be necessary for people who work in relatively higher dense areas where the risk of COVID-19 is high.

The counties in both the low and the median class of prevalence were accounted for 97.44% of the infected counties, where the higher values in the percentage of female population increased the mortality of COVID-19.

The percentage of adults with inadequate sleeping time was found to increase the mortality of COVID-19 in the low prevalence class of counties. Sleeping time was reported to be associated with the health system[25]. The higher number of people who had inadequate sleeping time, the more adverse effects of sleep on immunity were identified[26]. The air quality also was reported to be associated with the mortality rate of COVID-19[10, 30, 31].

For both the median and the high prevalence class of counties, there was an age trend in the mortality rate of COVID-19. In those counties, there was a higher percentage of elderly, indicating a larger population of individuals aged over 65, which increased the mortality rate of COVID-19[11].

Conclusion

This study identified several significant risk factors associated with the mortality of COVID-19, and our findings are highly valuable and timely for the decision-

makers to develop strategies in reducing the mortality of COVID-19. The study relied on mortality data on August 27, 2020. The counties were randomly divided into the training and testing data once. However, we offered the epidemiological picture to facilitate the identification of important factors influencing the mortality of COVID-19 across different levels of infected counties in the United States. Regardless of the regions, the factors linked to the poor health status contributed to higher mortality of COVID-19. Improving the clinical care and eliminating the racial health inequality, combined with improving physical environment were expected to significantly decrease the mortality rate of COVID-19. Thus, we recommended that local governments should reduce physical and psychological risks in residential environments.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The data that support the findings of this study are available from the New York Times and the County Health Rankings and Roadmaps program website. The data

and R files supporting the conclusions of this article are available in the https://github.com/tingT0929/Risk-factors-associated-with-mortality-of-COVID-19.

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Authors' contributions

TT, XW and HZ developed the idea and research. TT, JZ, LH and YJ wrote the first draft of the manuscript and all other authors discussed results and edited the manuscript. JZ, LH, and YJ collected and validated epidemiological data. CD and ZL collected and compiled census data. All authors read and approved the final manuscript

Acknowledgements

We would like to thank all individuals who are collecting epidemiological data of the COVID-19 outbreak around the world.

References

1. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J: **The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application**. *Annals*

of Internal Medicine 2020.

2. **Distribution of COVID-19 cases in the world** [http://2019ncov.chinacdc.cn/2019-nCoV/global.html]

3. **Coronavirus Disease 2019 (COVID-19)** [https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html]

4. Hadler JL, Yousey-Hindes K, Pérez A, Anderson EJ, Bargsten M, Bohm SR, Hill M, Hogan B, Laidler M, Lindegren ML: **Influenza-related hospitalizations and poverty levels—United States, 2010–2012**. *Morbidity Mortality Weekly Report* 2016, **65**(5):101-105.

5. Noppert GA, Yang Z, Clarke P, Ye W, Davidson P, Wilson ML: **Individual-and neighborhood-level contextual factors are associated with Mycobacterium tuberculosis transmission: genotypic clustering of cases in Michigan, 2004–2012**. *Annals of epidemiology* 2017, **27**(6):371-376. e375.

6. Maier HE, Lopez R, Sanchez N, Ng S, Gresh L, Ojeda S, Burger-Calderon R, Kuan G, Harris E, Balmaseda A: **Obesity increases the duration of influenza a virus shedding in adults**. *The Journal of infectious diseases* 2018, **218**(9):1378-1382.

7. Lin K, Fong DY-T, Zhu B, Karlberg J: **Environmental factors on the SARS epidemic: air temperature, passage of time and multiplicative effect of hospital infection**. *Epidemiology & Infection* 2006, **134**(2):223-230.

8. **State Rankings Data & Reports** [https://www.countyhealthrankings.org/reports/county-health-rankings-reports]

9. **Coronavirus in the U.S.: Latest Map and Case Count** [https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html]

10. Conticini E, Frediani B, Caro D: **Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy?** *Environmental pollution* 2020:114465.

11. Onder G, Rezza G, Brusaferro S: **Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy**. *JAMA* 2020.

12. Zhang LS, Yang MJ, Lei DJ: **An improved PAM clustering algorithm based on initial clustering centers**. In: *Applied Mechanics and Materials: 2012*: Trans Tech Publ; 2012: 244-249.

13. Lei D, Zhu Q, Chen J, Lin H: **Automatic PAM Clustering Algorithm for Outlier Detection**. *Journal of Software* 2012.

14. Brunner E, Konietschke F, Bathke AC, Pauly M: **Ranks and Pseudo-Ranks-Paradoxical Results of Rank Tests**. *arXiv preprint arXiv:180205650* 2018.

15. Hilbe JM: **Negative binomial regression**: Cambridge University Press; 2011.

16. Zeileis A, Kleiber C, Jackman S: **Regression models for count data in R**. *Journal of statistical software* 2008, **27**(8):1-25.

17. Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, Liu Y, Mills MC: **Demographic science aids in understanding the spread and fatality rates of COVID-19**. *Proceedings of the National Academy of Sciences* 2020.

18. Greenstone M, Nigam V: **Does social distancing matter?** *University of Chicago,*

Becker Friedman Institute for Economics Working Paper 2020(2020-26).

19. Hooper MW, Nápoles AM, Pérez-Stable E: **COVID-19 and Racial/Ethnic Disparities**. *JAMA* 2020.

20. Laurencin CT, McClinton Al: **The COVID-19 pandemic: a call to action to identify and address racial and ethnic disparities**. *Journal of Racial Ethnic Health Disparities* 2020:1-5.

21. Noonan AS, Velasco-Mondragon HE, Wagner FAJPhr: **Improving the health of African Americans in the USA: an overdue opportunity for social justice**. 2016, **37**(1):1-20.

22. Williams DR, Collins C: **Racial residential segregation**. *Race, Ethnicity, Health: A Public Health Reader*.2012, **26**:331.

23. Van Rooy DL: **Effects of automobile commute characteristics on affect and job candidate evaluations: A field experiment**. *Environment Behavior* 2006, **38**(5):626-655.

24. Wolin KY, Bennett GG, McNeill LH, Sorensen G, Emmons KMJAjohb: **Low discretionary time as a barrier to physical activity and intervention uptake**. 2008, **32**(6):563-569.

25. Besedovsky L, Lange T, Haack M: **The sleep-immune crosstalk in health and disease**. *Physiological reviews* 2019, **99**(3):1325-1380.

26. Irwin M: **Effects of sleep and sleep loss on immunity and cytokines**. *Brain, behavior, immunity* 2002, **16**(5):503-512.

27. Mazza C, Ricci E, Biondi S, Colasanti M, Ferracuti S, Napoli C, P. R: **A Nationwide Survey of Psychological Distress among Italian People during the COVID-19 Pandemic: Immediate Psychological Responses and Associated Factors**. *International Journal of Environmental Research and Public Health* 2020, **17**(9):3165.

28. Qiu J, Shen B, Zhao M, Wang Z, Xie B, Xu Y: **A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations**. *Gen Psychiatr* 2020, **33**(2):e100213-e100213.

29. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, Ho RC: **Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China**. *International journal of environmental research public health* 2020, **17**(5):1729.

30. Wu X, Nethery RC, Sabath BM, Braun D, Dominici F: **Exposure to air pollution and COVID-19 mortality in the United States**. *medRxiv* 2020.

31. Contini D, Costabile F: **Does Air Pollution Influence COVID-19 Outbreaks?** In.: Multidisciplinary Digital Publishing Institute; 2020.

32. Martelletti L, Martelletti P: **Air pollution and the novel Covid-19 disease: a putative disease risk factor**. *SN Comprehensive Clinical Medicine* 2020:1-5.

33. Wenham C, Smith J, Morgan R: **COVID-19: the gendered impacts of the outbreak**. *The Lancet* 2020, **395**(10227):846-848.

34. Abdulamir AS, Hafidh RR: **The Possible Immunological Pathways for the Variable Immunopathogenesis of COVID--19 Infections among Healthy Adults, Elderly and Children**. *Electronic Journal of General Medicine* 2020, **17**(4).

Table 1. The definitions and sources of the 12 selected risk factors, and the literature supporting these factors.

Category	Variable	Description	Source	Related Topics	Previous Studies
HEALTH BEHAVIORS	% With Access to Exercise Opportunities	Percentage of population with adequate access to locations for physical activity.	Business Analyst, Delorme map data, ESRI, & US Census Tigerline Files	Social Distancing; Outdoor and indoor exercise	Novel
	% Insufficient Sleep	Percentage of adults who report fewer than 7 hours of sleep on average.	Behavioral Risk Factor Surveillance System	Immune system	Novel
CLINICAL CARE	Primary Care Physicians Ratio	Ratio of population to primary care physicians.	Area Health Resource File/American Medical Association	Medicare services	Novel
SOCIAL & ECONOMIC FACTORS	Segregation index-non-Whites/Whites	Index of dissimilarity where higher values indicate greater residential segregation between non-White and White county residents.	American Community Survey, 5-year estimates	Poor health status	Novel
PHYSICAL ENVIRONMENT	% Long Commute - Drives Alone	Among workers who commute in their car alone, the percentage that commute more than 30 minutes.	American Community Survey, 5-year estimates	Psychological risks	Novel
	% Drive Alone to Work	Percentage of the workforce that drives alone to work.	American Community Survey, 5-year estimates	Psychological risks	Novel

	Average Daily PM2.5	Average daily density of fine particulate matter in micrograms per cubic meter (PM2.5).	Environmental Public Health Tracking Network	Air quality	E Conticini, B Frediani and D Caro [10] D Contini and F Costabile [31] L Martelletti and P Martelletti [32]
DEMOGRAPHICS	Population	Resident population.	Census Population Estimates	Population densities	Novel
	% Rural	Percentage of population living in a rural area.	Census Population Estimates	Demographical science	Novel
	% Hispanic	Percentage of population that is Hispanic.	Census Population Estimates	Minorities	Novel
	% Female	Percentage of population that is female.	Census Population Estimates	Sex	C Wenham, J Smith and R Morgan [33]
	% 65 and over	Percentage of population ages 65 and older.	Census Population Estimates	Age structure	G Onder, G Rezza and S Brusaferro [11], [34]

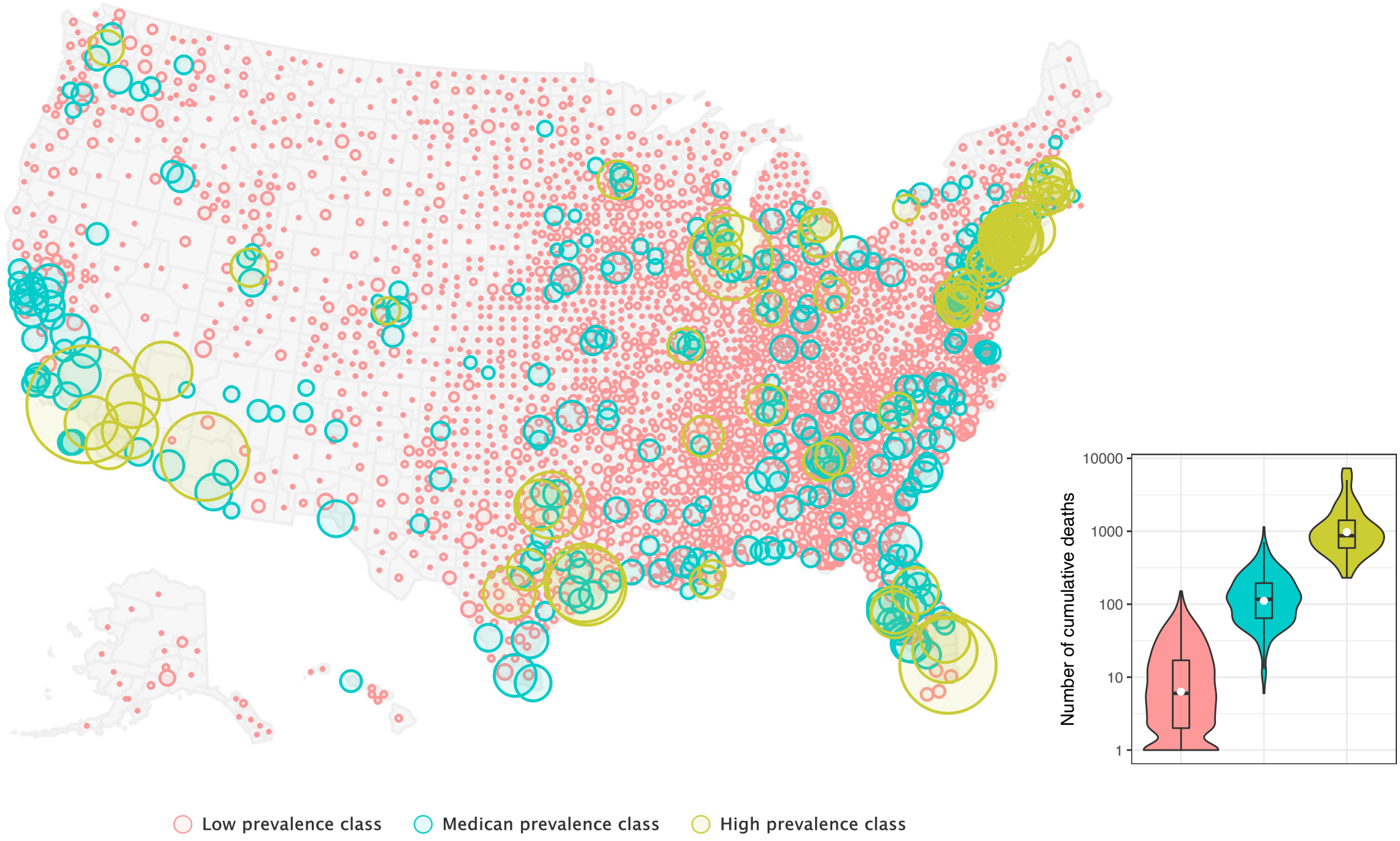
Table 2. Variables significantly related to the mortality rate of COVID-19 in the three classes of counties.

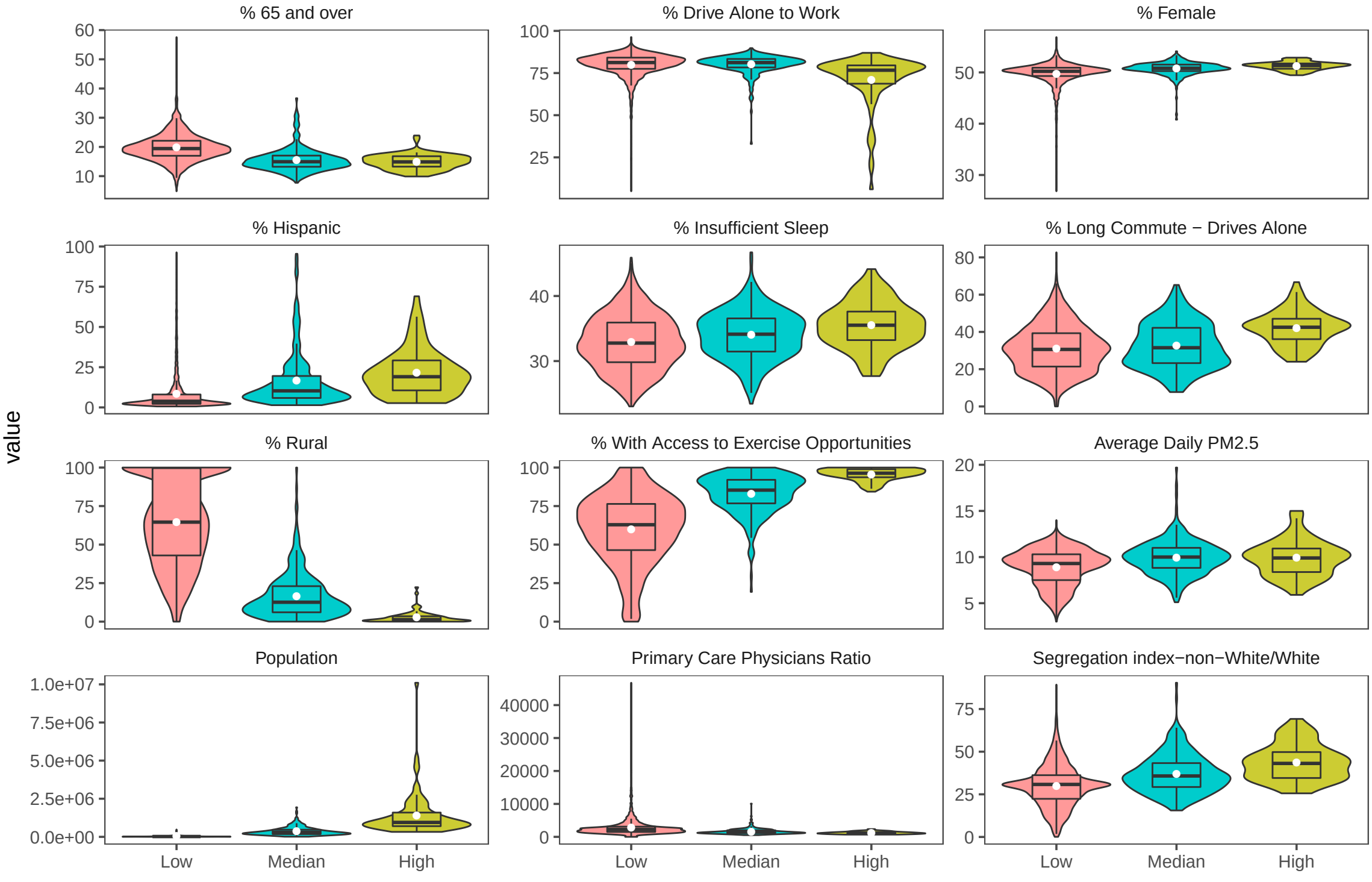
	Variable	Estimate	Pr(>|t|)
Low prevalence class	`Average Daily PM2.5`	0.153	<0.001
	`% Drive Alone to Work`	0.039	<0.001
	`% Long Commute - Drives Alone`	0.015	<0.001
	`% Insufficient Sleep`	0.073	<0.001
	Population	1.00E-05	<0.001
	`% Hispanic`	0.020	<0.001
	`% Female`	0.054	<0.001
	`% Rural`	-0.014	<0.001
	`Segregation index-non-White/White`	0.015	<0.001
Median prevalence class	`% Drive Alone to Work`	0.016	0.008
	`% Long Commute - Drives Alone`	0.013	<0.001
	Population	1.07E-06	<0.001
	`% 65 and over`	0.049	<0.001
	`% Hispanic`	0.020	<0.001
	% Female	0.127	0.001
	`% Rural`	-0.011	0.007
	`Segregation index-non-White/White`	0.025	<0.001
High prevalence class	Population	2.22E-07	<0.001
	`% 65 and over`	0.114	<0.001
	`%Hispanic `	0.014	0.009
	`Segregation index-non-White/White`	0.019	0.005

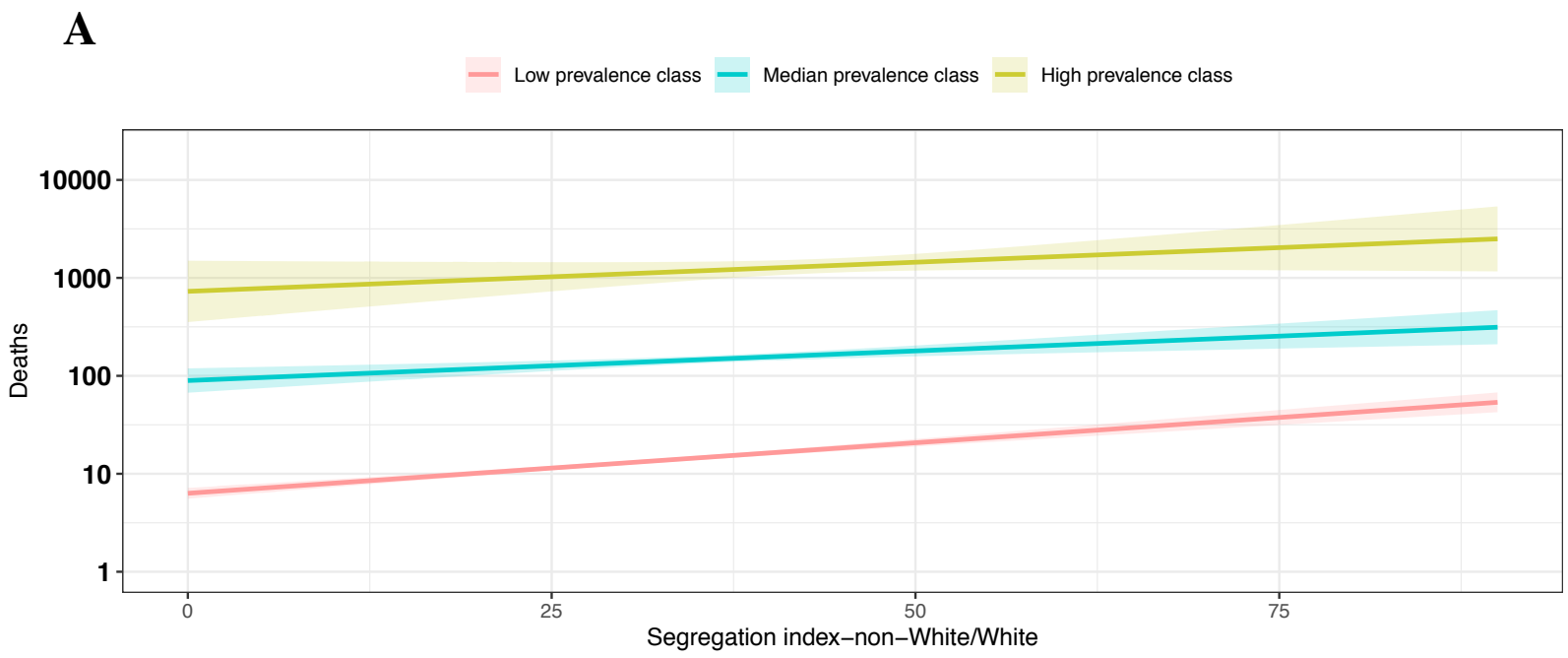
Figure 1. The geographical distribution of three classes of counties. The clustering was based on time-variant cumulative confirmed cases from March 1 to August 27, 2020. The size of circle represented the total confirmed cases on August 27, 2020. The distributions of deaths on August 27, 2020 in the three classes of counties were combined.

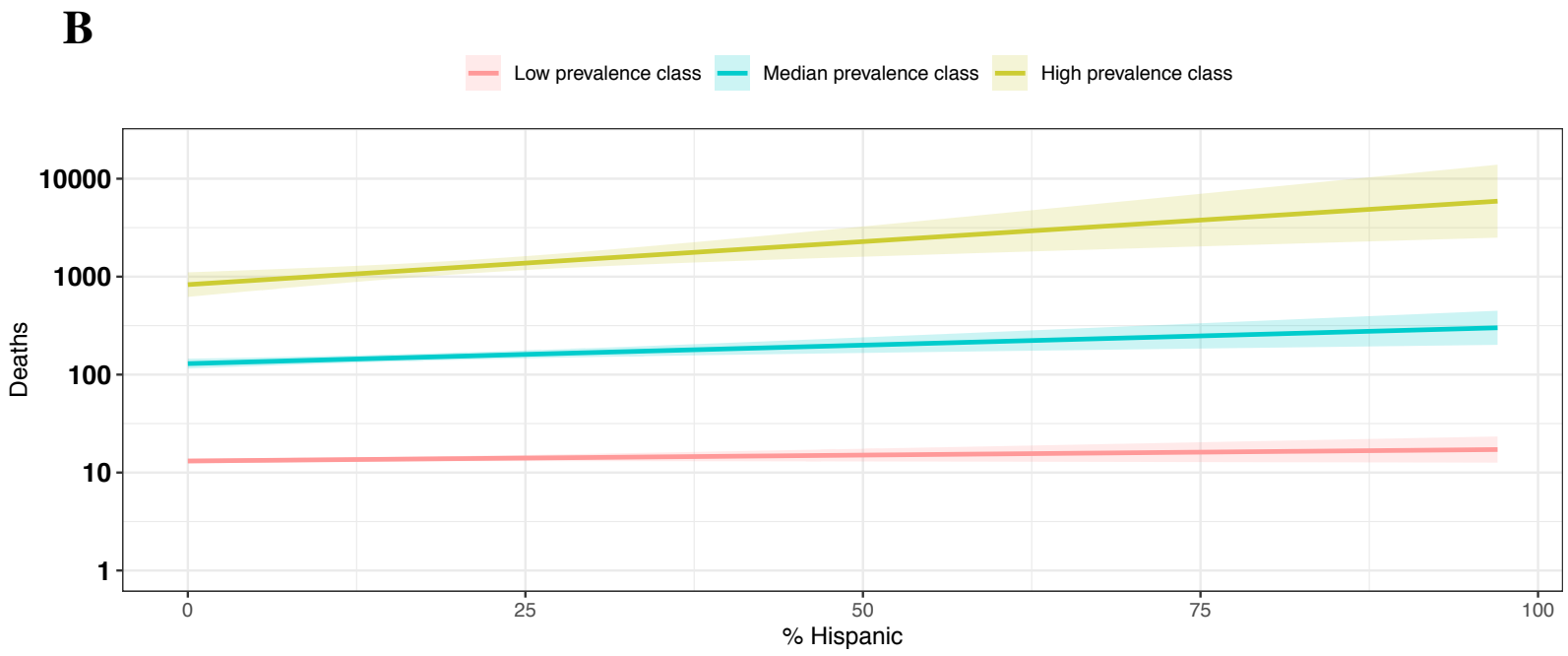
Figure 2. Violin diagram and boxplot of the distributions of the 12 selected risk factors in the three classes of counties. The low, median and high prevalence classes of counties were represented by red, blue and green colors.

Figure 3. The risk factors common in the three classes of counties. A: Segregation index for non-Whites versus Whites; B: The percentage of Hispanic population. Direct curves were generated using the negative binomial model with the segregation index between non-Whites and Whites or the percentage of Hispanic population as the single covariate.


Number of cumulative deaths
10000
1000
100
10
1
Low prevalence class
Medican prevalence class
High prevalence class


% 65 and over
% Drive Alone to Work
% Female
% Hispanic
% Insufficient Sleep
% Long Commute − Drives Alone
% Rural
% With Access to Exercise Opportunities
Average Daily PM2.5
Population
Primary Care Physicians Ratio
Segregation index−non−White/White
value
Low
Median
High


A
Low prevalence class
Median prevalence class
High prevalence class
Deaths
10000
1000
100
10
1
0
25
50
75
Segregation index−non−White/White


B
Low prevalence class
Median prevalence class
High prevalence class
Deaths
10000
1000
100
10
1
0
25
50
75
100
% Hispanic