Sociodemographic predictors and transportation patterns of COVID-19 infection and mortality

Renee Pekmezaris[1,2,3,4,5], Xu Zhu[1,3], Roland Hentz[4], Martin L. Lesser[1,2,3,4], Jason J. Wang[1,2,3,4], and Matthew Jelavic[6]

[1]Department of Medicine, Northwell Health, Manhasset NY 11549, USA
[2]Northwell Health, David and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead NY 11549, USA
[3]Insitute for Health Innovations and Outcomes Research, Northwell Health, Manhasset NY 11030, USA
[4]Northwell Health, The Feinstein Institute for Medical Research, Manhasset NY 11030, USA
[5]Division of Health Services Research, Center for Health Innovations and Outcomes Research, Department of Medicine, Manhasset NY 10030, USA
[6]Corporate Finance, Northwell Health, Westbury NY 11590, USA
Address correspondence to Xu Zhu, E-mail: xzhu2@northwell.edu

ABSTRACT

Background The United States Centers for Disease Control and Prevention (CDC)-sanctioned prevention strategies have included frequent handwashing with soap and water, covering the mouth and nose with a mask when around others, cleaning and disinfecting maintaining a distance of at least 6 feet from others, etc. Although many of these recommendations are based upon observation and past infection control practices, it is important to combine and explore public data sets to identify predictors of infection, morbidity and mortality to develop more finely honed interventions, based on sociodemographic factors.

Method Cross-sectional study of both states in the US and counties in NY state.

Results Population density was found to be significantly associated with state-level coronavirus infection and mortality rate ($b = 0.49$, 95% confidence interval (CI): 0.34, 0.64, $P < .0001$). States that have lower socioeconomic status, lower mean age and denser populations are associated with higher incidence rates. In regard to NY state, counties with a higher percentage of minority residents had higher COVID-19 mortality rates ($b = 2.61$, 95% CI: 0.36, 4.87, $P = 0.023$). Larger population cohorts were associated with lower COVID-19 mortality rates after adjusting for other variables in the model ($b = -1.39$, 95% CI: -2.07, -0.71, $P < 0.001$). Population density was not significantly associated with COVID-19 mortality rates after adjustment across counties in the NY state. Public ridership was not indicative of cases or mortality across states in the USA; however, it is a significant factor associated with incidence (but not mortality) in NY counties.

Conclusion Population density was the only significant predictor of mortality across states in the USA. Lower mean age, lower median household incomes and more densely populated states were at higher risk of COVID-19 infection. Population density was not found to be a significant independent variable compared to minority status and socioeconomic factors in the New York epicenter. Meanwhile, public ridership was found to be a significant factor associated with incidence in New York counties.

Keywords social determinants, morbidity and mortality, epidemiology

Population density and COVID-19

Since the unprecedented onset of the coronavirus disease 2019 (COVID-19) pandemic, multiple studies have fostered an understanding of correlates and predictors of infection and mortality in the literature. One such factor is population density—in theory, areas of higher population density facilitate more face-to-face interaction among residents, which can facilitate the rapid spread of pandemics. Conversely, areas of higher population density may have better access to health care facilities and more effective implementation of social

Renee Pekmezaris, Vice President, Professor
Xu Zhu, Research Scientist
Roland Hentz, Assistant Biostatistician
Martin L. Lesser, Associate Vice President
Jason J. Wang, Associate Professor
Matthew Jelavic, Vice President

distancing policies and practices. Several studies have researched population density and report conflicting results: one study from Wong and Li[1] found that population density is an effective predictor of cumulative infection cases in the USA on the county level from March to late May 2020. Meanwhile, a study of 913 metropolitan counties suggests that population density is not necessarily related to higher infection rates in the USA. Although larger metropolitan areas generally have higher infection and mortality rates, when controlling for metropolitan populations, county density is not significantly related to infection rates.[2] Carozzi[3] found that when controlling for other confounding factors such as geological features, population density is not associated with infection rate. Specifically, infection rates appear to have impacted denser areas earlier during the outbreak compared to sparser areas; however, this association did not hold after social distancing policies were implemented.[3]

To further investigate these mixed findings regarding population density, infection and death rates of COVID-19, the present study aimed to examine the impact of population density on both the national (state) and NY state (county) levels.

Public transportation and COVID-19

Since the emergence of COVID-19 in the USA in early 2020, transit ridership and fare revenues decreased significantly by April 2020 from April 2019—by 73% and 86%, respectively.[4] Specifically, New York City, one of the most affected regions in the state, restricted its public transportation policy in March of 2020. After implementation of the stay-at-home policy in March 2020, over 90% of the subway's 5.5 million weekday riders abandoned the system. Many studies have demonstrated that social distancing policies can effectively reduce the risk of infection. The number of COVID-19 cases and deaths were positively related to proximity to, and passenger volume of, the nearest airport, and the number of deaths and cases were positively associated with the number of transportation stations such as airports and train stations.[5]

Racial and ethnic disparities and COVID-19

According to the United States Centers for Disease Control and Prevention (CDC), although most infected cases are among non-Hispanic Whites (56.3%), racial minority groups are disproportionately represented among COVID-19 cases.[6] US census data form illustrates that Blacks account for 18.7% of overall deaths, despite representing just 12.5% of the US population.[7] Descriptive analyses from Gold and

colleagues[8] suggest that the percentage of Hispanic decedents infected with COVID-19 increased from 16.3% to 26.4% during May to August 2020. Indeed, racial and ethnic disparities have continued to persist during the course of COVID-19. This disparity seems to have worsened over time in non-metropolitan areas. In rural communities, minority status and language domains (persons for whom English is not their first language) persist as a driver of COVID-19 cases. This may, in part, be due to the fact that persons from underrepresented racial and ethnic minority groups are more likely to live in multigenerational and multifamily households, reside in congregate living environments, hold jobs requiring in-person work (e.g. meatpacking, agriculture, service and health care), have limited access to health care, or experience discrimination.[9,10] Further, greater social vulnerability has been shown to be associated with increased risk of COVID-19 detection and death.[10–12]

Social vulnerability index

The Social Vulnerability Index (SVI) from the CDC's US Community Census Survey is a measure of the socioeconomic status of a community. It was designed by the Geospatial Research, Analysis, and Services Program (GRASP) at the CDC. This index is often used to identify and map communities that will most likely need support before, during, and after a hazardous event. Figure 1 presents variables and themes included in the SVI databases that represent the components of SVI.[13]

Previous studies using the CDC's social vulnerability index identified associations between socioeconomic status and COVID-19 diagnosis and death. Khazanchi and Beiter's paper demonstrated that among rural counties alone, those with higher rates of minority status and language were associated with a greater risk of COVID-19 diagnosis.[14] Moreover, residents of the most vulnerable counties were found to have a greater risk of COVID-19 positive diagnosis and death by socioeconomic status (relative risks (RR) of 1.42 and 1.71) and housing and transportation (RR 1.52 and 1.32). A recent study using CDC's Social Vulnerability Index in combination with COVID-19 statistics suggests that the ability to stay home during the pandemic was constrained by socioeconomic status and work circumstances in the New York City area.[15] Increased subway use was also associated with a higher rate of COVID-19 cases per 100 000 population, when adjusted for testing effort but the association was weaker once adjusted for median income. Areas with lower median incomes and those with higher percentages of persons identifying as nonwhite had greater rates of subway use during the pandemic. This study also suggested that poorer neighborhoods are not as
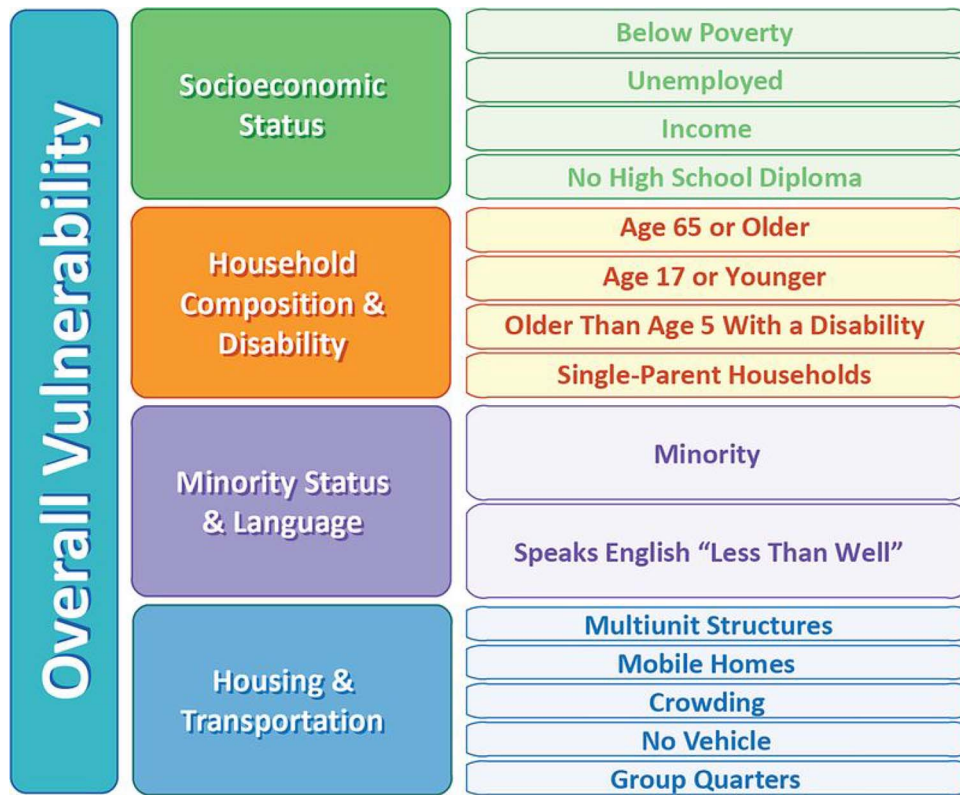
Fig. 1 Themes and variables in SVI.

likely to have the choice of mobility to move as their richer counterparts financially.

The present study was designed to explore the relationship between important sociodemographic variables and transportation/ridership data upon risk of infection and mortality. We hypothesize that:

(1) Areas with higher levels of public transportation ridership have significantly higher rates of COVID-19 cases and COVID-19-related mortality in the USA and NY.
(2) Areas with lower socioeconomic status, housing and transportation status (residents living in crowded, multi-unit housing and living without vehicles) are more likely to have higher rates of COVID-19 infection and mortality.
(3) Areas with higher levels of social vulnerability (e.g. the CDC's Social Vulnerability Index (SVI)) are more likely to have higher rates of COVID-19 infection and mortality; and
(4) Areas with denser populations are more likely to have higher rates of COVID-19 infection and mortality.

Of particular interest are the interactions between these variables and how they may help researchers to more finely predict COVID-19 infection and mortality.

Methods

The primary analyses of the impact of public transportation on COVID-19's spread and mortality were performed on two levels of analysis, state-level in the USA and county level in NY State, using cross-sectional analyses compiling various data sources from the New York Times, the CDC and American Transportation Association. State-level public transportation data were garnered from the American Transportation Association, reflecting the first Quarter of 2020. State-level Census data were extracted from US Census Data's Community Survey. We obtained COVID infection and mortality data from 1 March 2020 to 19 August 2020, as reported by the New York Times (https://github.com/nytimes/covid-19-data) and CDC's COVID data tracker (https://covid.cdc.gov/covid-data-tracker/).

Data sources

The datasets analyzed during the current study were extracted from the following public repositories:

(1) United States State level transportation data: to reflect weekly ridership by state, we used American Transportation Association data (https://www.apta.com/wp-content/uploads/2020-Q1-Ridership-APTA.pdf) from

2019 and 2020. Ridership is comprised of: Heavy rail (e.g. AMTRAK), light rail (subways), bus and trolley transportation. We incorporated all average weekly statistics from both sources.

(2) New York State County level data: demographic variables such as age and population reflecting the 3142 counties in the United States were extracted from the release in Dec 2019 release of 5-year estimates from US Census Bureau (2014–2018, https://www.census.gov/newsroom/press-releases/2019/acs-5-year.html)

(3) COVID-19 statistics: cases number and morbidity data were extracted from the CDC's COVID-19 data tracker (https://covid.cdc.gov/covid-data-tracker/#deaths) and data set is directed extracted from the New York Times (https://github.com/nytimes/covid-19-data).

(4) CDC SVI: the SVI from the CDC's US Community Census Survey is a representation of the socioeconomic status of a community. In addition to the overall measure, we explored the four SVI domains, including (i) the socioeconomic status domain; (ii) the household composition and disability domain; (iii) the minority status and language domain; and (iv) the housing type and transportation domain. We constructed the index at census tract level, a geographic scale commonly used to analyze community data for policy and planning in government and public health.[16] For the part two analyses, we also contracted databases at county level in NY State.

Statistical methods

Descriptive statistics were presented for all exposures and outcomes. Some variables were transformed using the Log2 function (Table 1). In the state-level analysis, the association between exposures and outcomes were tested using linear regression. Variables were screened at an inclusion criterion of $P < 0.1$, using univariable linear regression. Multivariable linear regressions were selected using backward elimination with a stay-criterion of $P < 0.1$ in the state-level analysis. In the county-level analysis, the association between exposures and outcomes were tested using robust regression. The MM-estimator method was used. Variables were screened at an inclusion criterion of $P < 0.1$ using univariable robust regression. Multivariable robust regression was selected using backward elimination with a stay criterion of $P < 0.1$. Significance for all tests was determined at $P < 0.05$.

Results

Part I: USA—state level.

Forty-seven states and Washington D.C. were used in the analysis. The observations of three states (MS, NE and NH)

were removed due to the absence of public transportation data. Detailed descriptive statistics are presented in Table 1.

Mortality rate

Univariable linear regression was applied to test our hypotheses regarding ridership and population density on mortality rates. States with both higher populations and population densities were significantly associated with higher COVID-19 mortality rates ($b = 0.41$, 95% confidence interval (CI): 0.16, 0.67, $P = 0.002$, and $b = 0.49$, 95% CI: 0.34, 0.64, $P < 0.001$, respectively). In addition, states with higher average weekly public transportation ridership were associated with higher COVID-19 mortality rates ($b = 0.21$, 95% CI: 0.12, 0.31, $P < 0.001$). The overall SVI was not associated with mortality; however, higher Minority Status and Language rankings (indicating higher percentages of minorities) were significantly associated with higher COVID-19 mortality rates ($b = 2.33$, 95% CI: 0.14, 4.52, $P = 0.037$) and higher Housing Type and Transportation rankings (representing residents who live in more crowded housing and own no vehicles) were associated with lower COVID-19 mortality rates ($b = -5.08$, 95% CI: -9.23, -0.93, $P = 0.018$). Additional variables that were found not to be associated with mortality included: mean age, states' SES rankings, state's Household Composition and Disability rankings.

The final multivariable model of mortality, selected by backward elimination, contained only a term for 'population density'. The estimates are equivalent to those given in the univariable results. Population density explained 48.7% of the variance of mortality rate in the model.

Incidence rate

The same predictors were tested on their univariable associations with incidence counts. States with higher populations and population densities were significantly associated with higher COVID-19 incidence rates ($b = 0.25$, 95% CI: 0.10, 0.41, $P = 0.002$, and ($b = 0.17$, 95% CI: 0.06, 0.29, $P = 0.004$, respectively). States with higher public transportation ridership were also associated with higher COVID-19 incidence rates ($b = 0.08$, 95% CI: 0.02, 0.15, $P = 0.013$). Among social vulnerability components, socioeconomic status and minority status 'statewide' are both significant factors predicting incidence rates, with states with a higher percentage of low household family incomes (including below poverty) having higher COVID-19 incidence rates ($b = 2.05$, 95% CI: 0.11, 3.99 $P = 0.039$) and higher percentage of minority tend to have higher COVID-19 rates ($b = 1.82$, 95% CI: 0.55, 3.09, $P = 0.006$; Table 3).

Table 1 Descriptive statistics of variables included in state-level analyses

Nationwide demographic information	Mean	SD	Median	IQR
N = 48				
Log2 mortality rate*	4.65	1.44	4.79	(3.80, 5.76)
Log2 incidence rate*	10.06	0.87	10.16	(9.77,10.67)
Log2 population	21.95	1.53	22.19	(20.77,22.94)
Log2 weekly public ridership*	13.48	3.77	14.25	(10.64,16.23)
Mean age	38.52	2.39	38.55	(37.15,39.60)
Log2 population density	6.62	2.06	6.74	(5.57,7.91)
Socio-economic ranking	0.47	0.13	0.47	(0.35,0.57)
Household composition ranking	0.52	0.13	0.53	(0.43,0.63)
Minority status ranking	0.41	0.19	0.42	(0.27,0.52)
Housing and transportation ranking	0.50	0.10	0.49	(0.43,0.55)
SVI ranking	0.47	0.11	0.47	(0.37,0.55)

*Detailed name of variables: log2 of mortality from COVID-19 per 100 000 population; log2 of incidence of COVID-19 per 100 000 population; average weekly public transportation ridership in 1000s; log2 of population density (pop./sq. mi.) per 100 000 population. SD, standard deviation.

After backward elimination, our multivariable model contained terms for 'mean age', 'population density' and 'SES ranking'. The estimates for these three variables were adjusted for each of the others. Higher population density was associated with higher COVID-19 incidence rates ($b = 0.24$, 95% CI: 0.14, 0.34, $P < 0.001$). Higher SES rankings (higher level of poverty) were associated with higher COVID-19 incidence rates ($b = 2.03$, 95% CI: 0.48, 3.58, $P = 0.012$). A higher mean age was associated with lower COVID-19 incidence rates ($b = -0.17$, 95% CI: -0.26, -0.08, $P < 0.001$). The model R^2 statistic was 0.443.

Part II: NY State—county level

Descriptive analyses showed 62 NY counties used in the analysis. The median log2 mortality rate per 100 000 population was 4.67 (interquartile range (IQR): 2.70, 6.01). The median log2 incidence rate per 100 000 population was 8.45 (IQR: 7.47, 9.65). Ridership across the counties was represented from the American Community Survey (ACS) question: 'percentage of employees who are older than 16 years old that use public transportation'. 'Total tests' referred to the combination of both diagnostic (polymerase chain reaction) tests and serologic tests from COVID-19 data source (Table 2).

Mortality rate

Univariable analysis was conducted to determine associations of mortality rates across counties in NY. Counties with a higher percentage of public ridership were associated with higher COVID-19 mortality rates ($b = 0.67$, 95% CI: 0.44, 0.89, $P < 0.001$). Counties with higher populations were associated with higher COVID-19 mortality rates ($b = 0.78$,

95% CI: 0.51, 1.05, $P < 0.001$) and so is population density ($b = 0.53$, 95% CI: 0.36, 0.71, $P < 0.001$). Moreover, total tests were associated with higher COVID-19 mortality rates ($b = 0.77$, 95% CI: 0.53, 1.01, $P < 0.001$), as were counties with higher Minority Status and Language rankings ($b = 5.26$, 95% CI: 3.76, 6.75, $P < 0.001$). Finally, counties with larger households and disability rankings were associated with lower COVID-19 mortality rates ($b = -2.19$, 95%: -4.36, -0.03, $P = 0.047$). SVI variables that were found not to be associated with mortality included: housing and transportation type, socio-economic ranking and SVI ranking.

Multivariable robust regression selected by backward elimination was adopted to test our hypothesis. The final model included the variables 'population density', 'total tests', 'population' and 'minority status ranking'. Counties with a higher percentage of minority residents had higher COVID-19 mortality rates ($b = 2.61$, 95% CI: 0.36, 4.87, $P = 0.023$). A higher total number of tests was also associated with higher COVID-19 mortality rates ($b = 1.21$, 95% CI: 0.74, 1.68, $P < 0.001$). Interestingly, larger population cohorts were associated with lower COVID-19 mortality rates after adjusting for other variables in the model ($b = -1.39$, 95% CI: -2.07, -0.71, $P < 0.001$). Population density was not significantly associated with COVID-19 mortality rates after adjustment. The final model indicated that these four factors (population density, total tests, population and minority ranking) accounted for 49% of variance in mortality rates in NY state.

Incidence rate

Univariable analyses were used to test the associations between independent variables and incidence rates. Counties

Table 2 Descriptive statistics of variables included in county-level analyses in NY State

NY county demographic information	Mean	SD	Median	IQR
N = 62				
Log2 Mortality Rate*	4.49	2.34	4.67	(2.7,6.01)
Log2 Incidence Rate*	8.78	1.86	8.45	(7.47,9.65)
Log2 Public Transport Ridership*	0.78	2.09	0.00	(−0.74,1.93)
Log2 Population Density	7.74	2.78	6.84	(6.2,8.75)
Log2 Total Tests	14.33	2.05	13.70	(13.07,15.59)
Log2 Population	16.89	1.87	16.40	(15.59,17.81)
Socio-economic ranking	0.50	0.30	0.49	(0.25,0.75)
Household composition ranking	0.50	0.30	0.50	(0.25,0.75)
Minority Status Ranking	0.50	0.30	0.50	(0.25,0.75)
Housing and transportation Ranking	0.50	0.29	0.50	(0.25,0.74)
SVI Ranking	0.50	0.30	0.50	(0.25,0.75)

*Detailed variable names: log2 of mortality from COVID-19 per 100 000 population, log2 of incidence of COVID-19 per 100 000 population, log2 of public transportation used over percent employees over 16 years of age.

Table 3 State-Level Results

State-level regression result estimates—mortality rate		
Variable	Estimate [95% CI]	P value
Log2 population density	0.49 [0.34–0.64]	<0.0001
State-level regression result estimates—incidence rate		
Mean age	−0.17 [0.26-0.08]	0.0004
Log2 population density	0.24 [0.14–0.34]	<0.0001
Socioeconomic ranking	2.03 [0.48–3.58]	0.0115

Table 4 New York County Results

NY county regression result estimates—mortality rate		
Variable	Estimate [95% CI]	P value
Log2 population density	0.34 [−0.03:0.70]	0.0745
Log2 total tests	1.21 [0.74: 1.68]	<.0001
Log2 population	−1.39[−2.07: −0.71]	<.0001
Minority language ranking	2.61 [0.36: 4.87]	0.0233
NY county regression result estimates—incidence rate		
Log2 Public Ridership	0.30[0.13:0.47]	0.0006
Log2 Total Tests	1.54[1.28: 1.81]	<.0001
Log2 Population	−1.31[−1.62: −0.99]	<.0001
Socioeconomic ranking	−1.33[−2.10: −0.56]	0.0007

with higher public transportation ridership were associated with higher COVID-19 incidence rates ($b = 0.60$, 95% CI: 0.46, 0.74, $P < 0.001$). Counties with higher population densities were also associated with higher COVID-19 incidence rates ($b = 0.45$, 95% CI: 0.34, 0.56, $P < 0.001$). Counties with larger populations were associated with higher

COVID-19 incidence rates ($b = 0.69$, 95% CI: 0.53, 0.85, $P < 0.001$). Numbers of total tests were also positively associated with incidence rates ($b = 0.66$, 95% CI: 0.52, 0.80, $P < 0.001$). Counties that had larger minority populations were estimated to have higher COVID-19 incidence rates ($b = 4.49$, 95% CI: 3.49, 5.49, $P < 0.001$). Surprisingly, higher socioeconomic rankings (indicating greater poverty) were associated with lower COVID-19 incidence rates ($b = -1.65$, 95% CI: -3.24, -0.05, $P = 0.044$). Surprisingly, counties with larger household composition and disability rankings were associated with lower COVID-19 incidence rates ($b = -2.50$, 95% CI: -3.99, -1.02, $P = 0.001$). Housing type and SVI rankings were not associated with COVID-19 incidence in NY state (Table 4).

The final model contained terms for 'public transportation ridership', 'total tests', 'population', and 'SES ranking'. The estimates for these four variables were adjusted for each of the others. As expected, higher public transportation ridership was associated with higher COVID-19 incidence rates ($b = 0.30$, 95% CI: 0.13, 0.47, $P < 0.001$). Higher numbers of total tests were also associated with higher COVID-19 incidence rates ($b = 1.55$, 95% CI: 1.28, 1.81, $P < 0.001$). Larger populations were again associated with lower COVID-19 incidence rates ($b = -1.31$, 95% CI: -1.62, -0.99, $P < 0.001$), as were higher SES rankings ($b = -1.33$, 95% CI: -2.10, -0.56, $P < 0.001$). The fit statistic for the model was $R^2 = 0.578$ (Table 4).

Discussion

Main findings of this study

Despite strong positive associations of weekly public ridership and minority population proportions on mortality and incidence rates, multiple regression using backward elimination suggests that population density is the only significant predictor of mortality across states in the USA. We also found that mean age, population density and socioeconomic status are significant predictors of 'incidence rates' across the USA, with lower mean age, lower median household incomes and more densely populated states at higher risk of COVID-19 infection.

Our findings in the NY State analysis are consistent with previous studies. Similar to Almagro et al's[9] findings, population density was not found to be a significant independent variable compared to minority status and socioeconomic factors in the NY epicenter. In addition, public ridership is not indicative of cases or mortality across states in the USA; however, it is a significant factor associated with incidence (but not mortality) in New York counties.

Our two levels of analysis—national and state—shows that the dynamics behind the spread and patterns of mortality and incidence across various demographic groups are highly complex.

What is already known on this topic

First, mixed findings on population density and it association with COVID-19 incidence were discovered through literature.

Second, we already know minorities in the USA have been disproportionately infected and this disparity seems to have worsened over time in non-metropolitan areas.

What this study adds

This study is different from previous investigations in several ways. First, our ridership data, manually extracted from the APTA website, represents average weekly ridership, reflecting transportation options in each state. Ridership nationwide is comprised of heavy rails, light rails, buses and trolley statistics. No study had used that data. Second, both incidence and mortality rates were used as dependent variables, as they are both important indicators of COVID-19 prevalence. National (state) data and New York county data were examined in relation to both incidence and mortality rates. Many previous studies looked only at mortality rates.

Moreover, our analysis of population density and public ridership sheds light on their effects on mortality and incidence during the current COVID-19 pandemic in the USA. The finding that frequency use of public transportation is significant only for the New York data set is, at least in part, due to the geographic and sociodemographic uniqueness of New York City. Specifically, the New York metropolitan area represents a large number of people living and working in a relatively small geographic area bordered by water on three sides, reflecting the 'interconnectedness' of persons in a closed environment, which increases the risk of exposure of COVID-19.

In addition, it is interesting that states with a higher mean age are more likely to have lower infection rates; this may in part be due to increased precautions taken by older adults, or, conversely, decreased precautions taken by their younger counterpart.

Limitations of this study

In our NY state incidence analyses, population density was not found to be statistically significant with regard to mortality even though the opposite was true in the univariable model. How and why this trend reverses in the multivariable model is important to understand in future investigations. As this study is cross-sectional in nature, thus future studies should develop longitudinal, time series analyses to identify the population density and transportation impact trends.

Acknowledgements

Funding

Conflict of interest

None.

Presentation

No presentation was made previously.

References

1 Wong DWS, Li Y. Spreading of COVID-19: density matters. *PLoS One* 2020;**15**(**12**):e0242398.

2 Hamidi S, Sabouri S, Ewing R. Does density aggravate the COVID-19 pandemic?: Early findings and lessons for planners. *J Am Plann Assoc* 2020;**86**(**4**):1–15.

3 Carozzi, F, Provenzano S, Roth S, *Urban Density and Covid-19*. IZA Discussion Paper, UK, No. 13440. 2020

4 APTA. *Impact of COVID-19 Pandemic on Public Transit Funding Needs*, United States, 5 May 2020.

5 Zhang Y, Fricker JD. Quantifying the impact of COVID-19 on non-motorized transportation: A Bayesian structural time series model. *Transport Policy* 2021;**103**:11–20ISSN 0967-070X. doi: 10.1016/j.tranpol.2021.01.013.

6 US Census Bureau. *Quick Facts*. Suitland, MD: US Department of Commerce, US Census Bureau, 2020.

7 Gold JA, Rossen LM, Ahmad FB *et al.* Race, Ethnicity, and Age Trends in Persons Who Died from COVID-19 — United States, May–August 2020. *MMWR Morb Mortal Wkly Rep* 2020;**69**(**42**): 1517–21.

8 New York State Department of Health. *COVID-19 fatalities*. Updated 11 April 2020.. https://covid19tracker.health.ny.gov/views/NYS-COVID19-Tracker/ (12 April 2020, date last accessed).

9 Almagro M, Orane-Hutchinson A. The determinants of the differential exposure to COVID-19 in New York City and their evolution over time. *Covid Economics: Vetted and Real-Time Papers* 2020;**13**.

10 Gaskin DJ, Zare H, Delarmente B. Geographic disparities in COVID-19 infections and deaths: the role of transportation. *Transport Pol* 2020. doi: 10.1016/j.tranpol.2020.

11 Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. *JAMA* 2020;**323**(**24**):2466–7.

12 Bui DP, McCaffrey K, Friedrichs M *et al.* Racial and ethnic disparities among COVID-19 cases in workplace outbreaks by industry sector—Utah, March 6–June 5, 2020. *MMWR Morb Mortal Wkly Rep* 2020;**69**(**33**):1133.

13 Flanagan BE, Gregory EW, Hallisey EJ *et al.* A Social Vulnerability Index for Disaster Management. *J Homeland Security Emergency Manag* 2011;**8**(**1**):0000102202154773551792. https://doi.org/10.2202/1547-7355.1792.

14 Khazanchi R, Beiter ER, Gondi S *et al.* County-level association of social vulnerability with COVID-19 cases and deaths in the USA. *J Gen Intern Med* 2020;**35**(**9**):2784–87.

15 Sy KTL, Martinez ME, Rader B, White LF. Socioeconomic disparities in subway use and COVID-19 outcomes in New York City. medRxiv [Preprint]. 2020 May;**30**:2020.05.28.20115949. United States. doi: 10.1101/2020.05.28.20115949.

16 Krieger N. A century of census tracts: Health & the body politic (1906–2006). *J Urban Health* 2006;**83**(**3**):355–61.