THE JOURNAL OF RURAL HEALTH

ORIGINAL ARTICLE

Spatial Disparities in Coronavirus Incidence and Mortality in the United States: An Ecological Analysis as of May 2020

Charlie H. Zhang, PhD [1] & Gary G. Schwartz, PhD, MPH, PhD[2]

1 Department of Geography & Geosciences, University of Louisville, Louisville, Kentucky
2 Department of Population Health, School of Medicine & Health Sciences, University of North Dakota, Grand Forks, North Dakota

Acknowledgments: The authors thank ESRI and Johns Hopkins University for releasing timely open access data on COVID-19.

For further information, contact: Charlie H. Zhang, PhD, Department of Geography & Geosciences, University of Louisville, 216 Lutz Hall, 2301 S. 3rd St., Louisville, KY 40292; e-mail: c.zhang@louisville.edu.

doi: 10.1111/jrh.12476

Abstract

Purpose: This ecological analysis investigates the spatial patterns of the COVID-19 epidemic in the United States in relation to socioeconomic variables that characterize US counties.

Methods: Data on confirmed cases and deaths from COVID-19 for 2,814 US counties were obtained from Johns Hopkins University. We used Geographic Information Systems (GIS) to map the spatial aspects of this pandemic and investigate the disparities between metropolitan and nonmetropolitan communities. Multiple regression models were used to explore the contextual risk factors of infections and death across US counties. We included population density, percent of population aged 65+, percent population in poverty, percent minority population, and percent of the uninsured as independent variables. A state-level measure of the percent of the population that has been tested for COVID-19 was used to control for the impact of testing.

Findings: The impact of COVID-19 in the United States has been extremely uneven. Although densely populated large cities and their surrounding metropolitan areas are hotspots of the pandemic, it is counterintuitive that incidence and mortality rates in some small cities and nonmetropolitan counties approximate those in epicenters such as New York City. Regression analyses support the hypotheses of positive correlations between COVID-19 incidence and mortality rates and socioeconomic factors including population density, proportions of elderly residents, poverty, and percent population tested.

Conclusions: Knowledge about the spatial aspects of the COVID-19 epidemic and its socioeconomic correlates can inform first responders and government efforts. Directives for social distancing and to "shelter-in-place" should continue to stem the spread of COVID-19.

Key words COVID-19, pandemic, metropolitan areas, rural or nonmetropolitan, spatial disparities.

COVID-19 is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) that originated December 2019 in Wuhan, China, a metropolis with a population of more than 11 million.[1] The outbreak proliferated rapidly worldwide and was declared a pandemic within months. At the time of this writing (May 1, 2020), SARS-COV-2 has infected approximately 3.5 million people across 187 countries (or regions) and has caused nearly 240,000 deaths. The first US case was reported on January 19, 2020, in Snohomish County, Washington.[2] Since then, the United States has become the epicenter of the pandemic with more than 1,000,000 confirmed cases and over 60,000 deaths. Epidemiological predictions suggest that, depending upon the location within the United States, it may take weeks for the growth curve of US cases to peak.

This is an open access article under the terms of the CreativeCommonsAttribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Transmission of SARS-COV-2 occurs primarily via the respiratory route whereby droplets containing the virus travel when an infected individual coughs or sneezes. It is unclear whether the virus also is transmitted via the aerosol route; that is, via smaller particles not requiring droplet transmission.[3,4] Early research found that most infected individuals suffer only minor symptoms.[5-7] Conversely, older individuals, particularly those with coexisting morbidities (eg, hypertension, diabetes, asthma, and/or obesity), are more likely to experience severe symptoms and to die.[8-10] Other data, albeit largely anecdotal, indicate that the death toll from infection is disproportionately greater among minorities and among residents who live in inner cities and/or rural communities.[11] The greater vulnerability of these groups is believed to be due to a combination of poverty, a higher prevalence of comorbid diseases, and poor accessibility to health care.[12] In terms of the differences between urban and rural areas in exposure to COVID-19, although the high population density in large cities and metropolitan areas is a major risk factor for contracting the virus, rural areas may be uniquely vulnerable due to the older age structure of many rural communities, the higher prevalence of chronic illnesses (eg, diabetes and hypertension), and a relative lack of health care facilities and services.[13-18]

Since the outbreak of SARS-COV-2 in the United States, Geographic Information Systems (GIS) have been used to quickly map the distribution and diffusion of the pandemic.[19-22] However, to our knowledge, there has been little published analysis of geographic disparities in infections and deaths among urban, suburban, and rural communities.[23,24] Additionally, empirical data that explore the risk factors that are associated with SARS-COV-2 infections and fatalities are needed in order to inform policies including preparation, response, mitigation, and recovery strategies.[25-27] This ecological analysis investigates spatial patterns of COVID-19 relative to socioeconomic contexts of different types of communities across the United States using data for US counties and metropolitan areas.

Data and Methods

Data on confirmed cases and deaths of COVID-19 were compiled and released by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, which compiled data collated by the US Centers for Disease Control & Prevention (CDC) from individual states and local health agencies. We used county-level data for 2,814 US counties that had at least one case of infection as of May 1 2020. The data for New York City includes its 5 boroughs (or counties): Manhattan, Bronx, Kings, Queens, and Richmond, which were provided as a single unit in the original data set. Numbers of cases and deaths for Kansas City, Missouri, and Jackson County (home to Kansas City) were reported separately in the original data but were consolidated as one county. We excluded cases that were assigned at the state level but were not affiliated with a specific county. We calculated incidence rates and death rates per 100,000 for each county and metropolitan area using the 2018 population data from the US Census Bureau.

In order to examine spatial disparities between metropolitan and nonmetropolitan counties, we used urban-rural continuum locale codes from the US Department of Agriculture (USDA).[28-30] Briefly, the USDA codes are based on the size of a county's urban population and its location relative to metropolitan areas of different sizes. Counties located inside metropolitan statistical areas (MSAs) are divided into 3 subcategories: counties in the largest MSAs (1 million population or more) are assigned a locality of 1; counties with populations between 250,000 and 1 million, a locality of 2; and those with populations smaller than 250,000, a locality of 3. Counties located outside any metropolitan area are assigned a subcategory ranging from 4 to 9 depending on the size of its urban population and whether it is adjacent to a metro area. Among the 2,814 counties (nearly 90% of all counties in the United States) that have reported data on COVID-19 infections and deaths, 431 are in large MSAs, 370 in medium-sized MSAs, 346 in small MSAs, and 1,667 are nonmetropolitan or rural counties. Additionally, the Census Bureau classifies metropolitan counties into 2 subcategories: central or urban counties and outlying or suburban counties. These subcategories help to differentiate urban and suburban communities within metropolitan areas.

To investigate the spatial variations of COVID-19 cases and deaths, we performed multiple regression analyses using a subset of the original data—1624 counties with 16 or more cases—as recommended by CDC, in order to mitigate the analytical problems caused by counties with a small number of cases and to preserve the anonymity of cases.[31] Confirmed cases and deaths for each county were used as the dependent variables, respectively. We used data on county-level social vulnerability indices in 2018 created by the CDC as independent variables. These indices included the percent of population aged 65+, the percent of the population in poverty, the percent minority, and percent uninsured. Population density was considered the primary predictor since more densely populated communities make it easier for infections to spread; indeed, this is the rationale for the public health intervention of "social distancing."[32] Because older individuals have a higher prevalence of preexisting chronic

Table 1 Top 10 Metropolitan Areas with Largest Confirmed Cases of COVID-19

Ranking	Metropolitan Areas	Population	Infections	Deaths
1	New York-Newark-Jersey City, NY-NJ-PA	19,990,592	394,259	29,373
2	Chicago-Naperville-Elgin, IL-IN-WI	9,536,428	50,705	2,257
3	Boston-Cambridge-Newton, MA-NH	4,811,732	47,583	2,624
4	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	6,069,448	36,508	1,852
5	Detroit-Warren-Dearborn, MI	4,317,179	30,344	3,145
6	Los Angeles-Long Beach-Anaheim, CA	13,262,234	25,626	1,164
7	Washington-Arlington-Alexandria, DC-VA-MD-WV	6,138,382	25,163	1,080
8	Miami-Fort Lauderdale-West Palm Beach, FL	6,070,944	19,979	723
9	New Orleans-Metairie, LA	1,263,635	16,094	1,053
10	Atlanta-Sandy Springs-Roswell, GA	5,779,463	13,211	505

illnesses and are more vulnerable to symptomatic infections, we included the percent of the population aged 65 and above.[33] In light of well-documented health disparities between rich and poor and between Whites and minorities (eg, Blacks, Hispanics, Native Americans), we hypothesized that the poverty level and the percentage of minority population would be positively correlated with disease burden.[34-37] However, an assessment of the number of infections and deaths of COVID-19 is largely dependent on the availability of testing.[38] We therefore included the percent of the population being tested at the state level, as testing data were not available at the county level.

Additionally, we conducted sensitivity analyses using a smaller sample that comprised only 688 rural counties and including population density as the only independent variable in regression analysis to explore whether population density is a helpful predictor or not. Because most rural areas are sparsely populated in the United States, it was expected that population density would be less effective in predicting the variations of coronavirus infections and deaths across rural communities.

Results

Disparities in COVID-19 Between Metropolitan and Nonmetropolitan Areas

We excluded 7,445 cases (less than 1% of the total) that were not affiliated with a specific county. The COVID-19 pandemic has affected 90% or 2,814 of US counties in all 50 states and Washington DC, among which large cities and their associated metropolitan areas have been affected the most. New York City, the nation's largest city (2018 population of 8.4 million in its 5 boroughs), recorded 167,478 confirmed cases, representing 15.7% of the nation's total infections. The New York metropolitan area, which comprises 25 counties that span New York, New Jersey, and Pennsylvania (2018 population of 20 million), accounts for nearly 37% (or 394,259 confirmed cases) of the nation's infections (Table 1). After the New York metro, other metropolitan areas, including Chicago, Boston, Philadelphia, Detroit, Los Angeles, Washington DC, Miami, New Orleans, and Atlanta, are among the top 10 in terms of number of confirmed incident cases. The top 10 metropolitan areas represent nearly 61% of all confirmed cases of COVID-19 in the United States.

Large metropolitan areas also have the highest number of coronavirus fatalities (Table 1). With 29,373 deaths, the New York MSA accounts for nearly 47% of the total deaths nationwide, followed by other MSAs including Detroit, Boston, Chicago, Philadelphia, Los Angeles, Washington DC, and New Orleans. It is noteworthy that Detroit and New Orleans have become the second and eighth largest hotspots despite their population rankings, which are considerably lower (the 14th and 24th largest MSAs). Los Angeles and Chicago, the second and third largest metropolitan areas, on the contrary, ranked lower in terms of both infections and deaths than their corresponding populations would suggest.

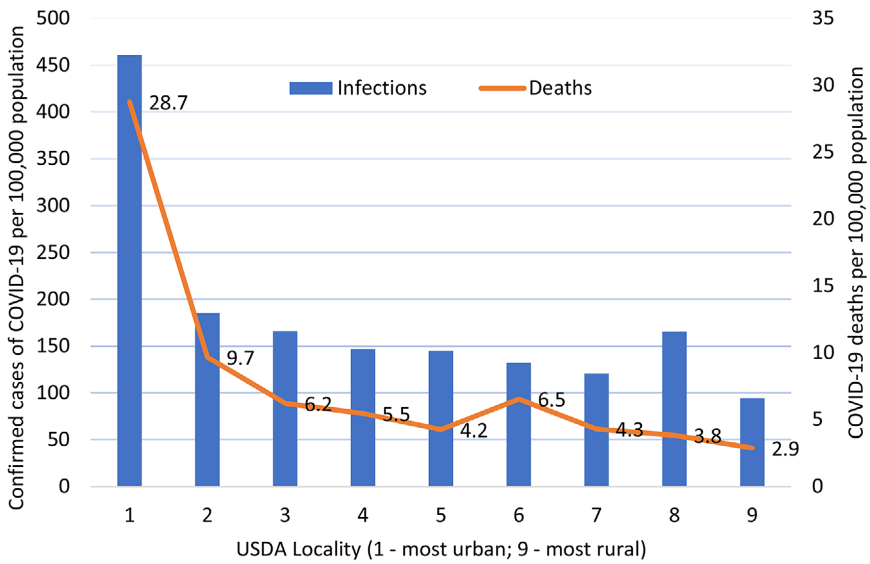
More than 94.3% of all confirmed cases (1,061,223 of 1,068,668 as of May 1) and 96.3% of all deaths (62,295 out of 62,698) were recorded in 1,447 metropolitan counties (USDA locale codes 1–3) (Table 2). In other words, 5.6% (59,940) of all infections and 3.7% (2333) of deaths were recorded in 1,667 nonmetropolitan or rural counties (USDA locale codes 4–9). Metropolitan counties, especially those located in large MSAs with populations more than 1 million, had the highest rates of infection and death rates. Rural or nonmetropolitan counties, in general, had the lowest rates for both infection and mortality, especially in nonmetro counties with populations less than 19,999 and not adjacent to any MSA (ie, locales 7 and 9) (Figure 1). However, it is noteworthy that some counties that are completely rural and/or those adjacent to MSAs but with urban populations less than 2,500 (ie, locale 8) had an average infection rate nearly as high as that in small MSA counties (ie, locale 3). Likewise, nonmetro counties with urban population between 2,500 and

Table 2 Confirmed Cases and Deaths of COVID-19 Pandemic Summarized by Urban-Rural Locality*

Locality	Counties	Infections	%	Deaths	%
1 – Metro counties in MSAs of 1 million population or more (Large MSA)	431	826,750	77.9	51,553	82.8
2 – Metro counties in MSAs of 250,000 to 1 million population (Medium-sized MSA)	370	126,612	11.9	6,616	10.6
3 – Metro counties in MSAs of fewer than 250,000 population (Small MSA)	346	47,921	4.5	1,793	2.9
4 - Nonmetro—Urban population≥ 20,000, adjacent to a MSA	213	19,879	1.9	738	1.2
5 - Nonmetro—Urban population≥ 20,000 or more, not adjacent to a MSA	92	7,305	0.7	214	0.3
6 - Nonmetro—Urban population of 2,500 to 19,999, adjacent to a MSA	565	18,805	1.8	932	1.5
7 - Nonmetro—Urban population of 2,500 to 19,999, not adjacent to a MSA	375	9,052	0.9	322	0.5
8 - Nonmetro—Completely rural or <2,500 urban population, adjacent to a MSA	170	3,066	0.3	71	0.1
9 - Nonmetro—Completely rural or <2,500 urban population, not adjacent to a MSA	252	1,833	0.2	56	0.1
Total number of counties	2,814 (1,147 metro; 1,667 nonmetro)	1,061,223	100	62,295	100

*A total of 7,445 (or less than 1%) of confirmed cases and 643 (or less than 1%) of deaths in the original data were excluded because they were assigned at the state level but were not affiliated with a specific county.

Figure 1 Average Incidence Rates of Confirmed Cases and Deaths from COVID-19 Pandemic for US Counties Classified by USDA Urban-Rural Locale Codes (2013).



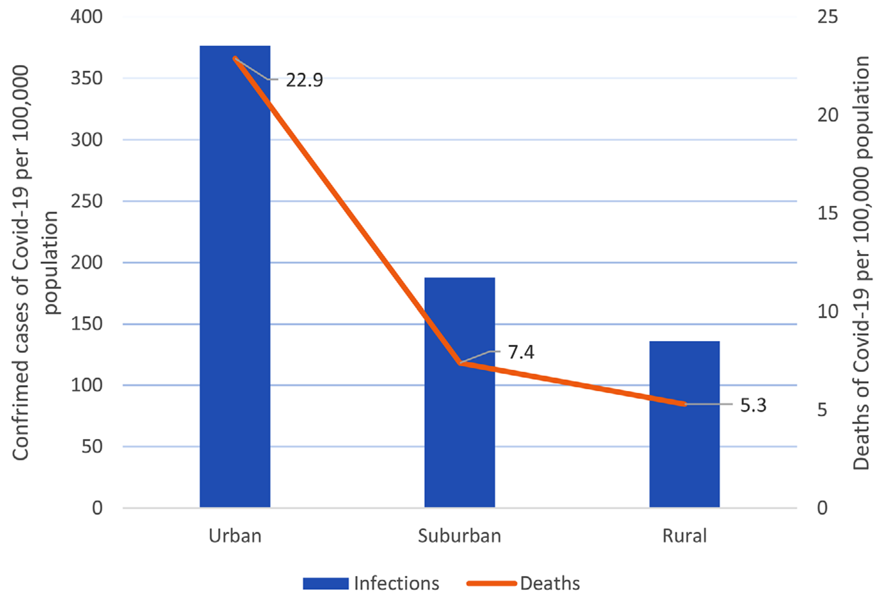
19,999 and adjacent to MSAs (ie, locale 6) had an average mortality rate just below large and medium MSA counties.

Within metropolitan areas, central or urban counties had much higher incidence and mortality rates than outlying or suburban counties (Figure 2). In summary, urban counties had the highest levels, rural counties the lowest levels, and suburban counties intermediate levels of incident and fatal COVID-19 cases.

Mapping Spatial Patterns of Covid-19 Infections and Deaths

Mapping incidence and mortality rates at a more granular scale, for example, individual counties that are the building blocks of metropolitan areas, allows us to visualize more detailed patterns of COVID-19 (Figures 3 and 4). As noted above, it is not surprising that the highest infection rates are in large cities and their associated metropolitan areas, especially the Northeast corridor that stretches from Boston to Washington DC, the Midwest (ie, Detroit and Chicago), and New Orleans. In addition to large metropolitan areas, small metropolitan areas and rural counties that are far away from major cities, including southwest Georgia, northeast Arizona, and northwest New Mexico, that spans a large portion of the Navajo Nation, west-central Colorado, and south-central Idaho, show surprisingly high rates of confirmed cases. An inspection of the top 25 counties with the highest

Figure 2 Average Incidence and Mortality Rates in Urban, Suburban, and Rural Counties in the United States.



infection rates shows that 3 counties are in small MSAs and 11 counties are rural or nonmetro (see Table 3). Lincoln County, Arkansas, an outlying county (population 13,695 in 2018) in the small Pine Bluff MSA, had the highest infection rate (ie, 5,943.8 per 100,000 population), nearly 3 times larger than that in New York City (ranked 22 with an incidence rate of 1983.5 per 100,000 population). Thus, several small metro areas and some rural areas are notable exceptions to the generalization that COVID-19 cases are limited to large metropolitan areas.

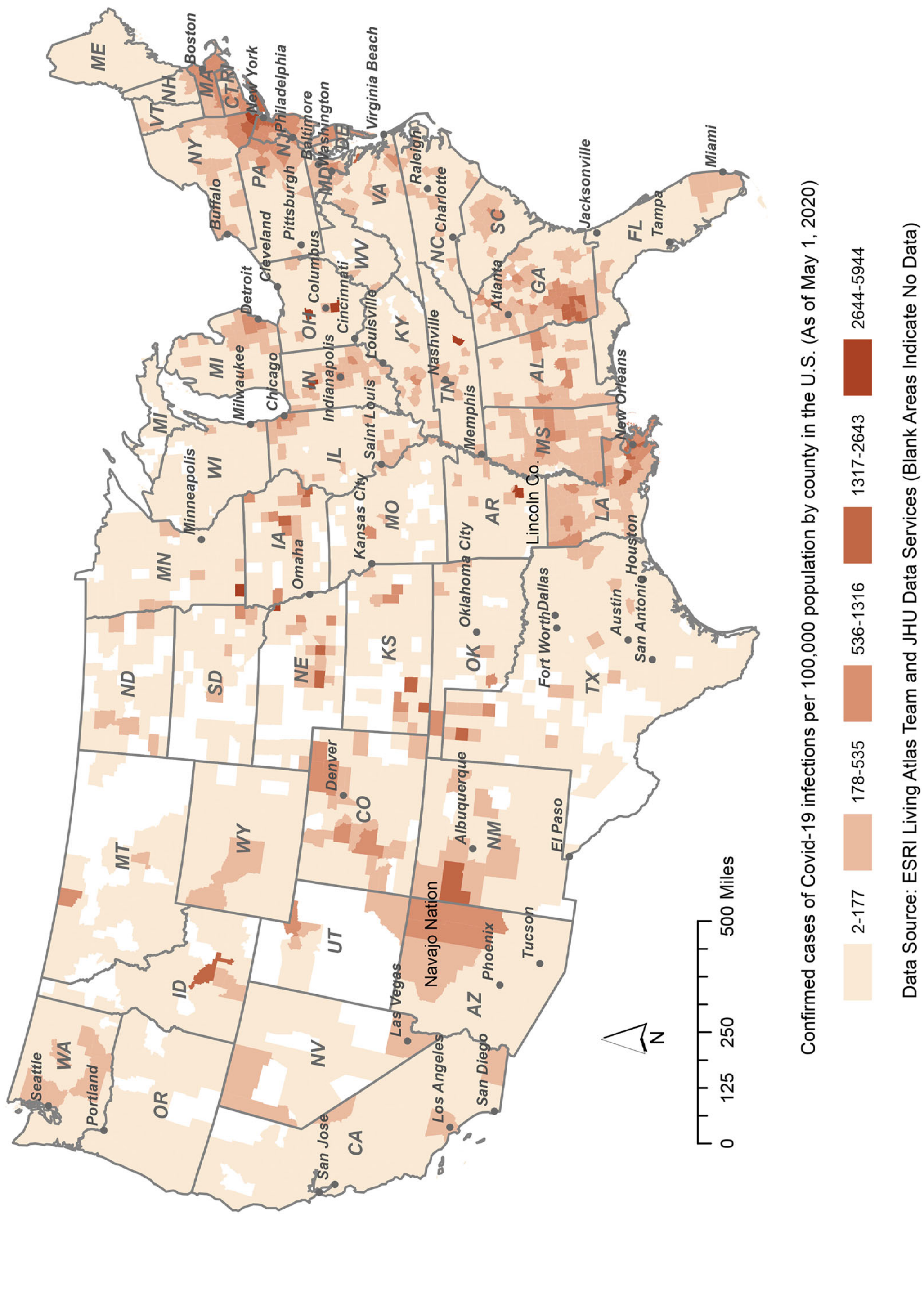
An examination of COVID-19 mortality rates underscores the observation that small metropolitan areas and some nonmetropolitan counties are disproportionately affected (Table 4). For example, the second-highest death rate in the nation (after the New York MSA) is in the Albany metropolis, a small MSA in southwest Georgia, population 153,101. Among the top 25 MSAs, 8 are small metropolitan areas with populations less than 250,000, including Albany, GA; Grand Island, NE; Houma-Thibodaux, LA; Farmington, NM; East Stroudsburg, PA; Lewiston, ID-WA; Saginaw, MI; and Flagstaff, AZ. MSAs of Albany, GA and Grand Island, NE also rank second and third in infection rate among all metropolitan areas.

The county with the highest mortality rate (268.1 deaths per 100,000) is Randolph, GA (population 7,087), a rural county near the Alabama border (Table 5). Among the top 25 in mortality rates, 2 Georgia counties, Dougherty and Terrell, are located in the Albany (GA) metropolitan area. Eleven (11) are classified as rural nonmetropolitan counties that are scattered in a number of states including Georgia (Clay, Early, Mitchell, Randolph, Turner, Wilcox counties), Indiana (Decatur County), Kansas (Coffey County), Louisiana (Bienville County), Montana (Toole county), and Oklahoma (Greer County).

Ecological Analysis of the Correlates of COVID-19

Results of regression models for COVID-19 incident cases are presented in Table 6. As expected, the variables population density, percent population aged 65+, and percent population tested for COVID-19 showed statistically significant and positive correlations with confirmed cases of COVID-19. Population density is the strongest predictor of the variations of infections among counties and explains 43% of the variations of infections. The impact of testing rates is understandable and substantiates epidemiologists' call for expanding testing to control the spread of coronavirus. Poverty and percent minority population were not significant and showed negative correlations with incidence rates. One possible explanation is that this reflects the likelihood that poverty and minority populations are factors that lead individuals to be untested. The percent uninsured population showed a positive sign but was not statistically significant, which could have also been affected by shortages of testing. Overall, the regression model explained nearly half of the variation in incidence rates (45%) of COVID-19 across 1,624 counties. Multicollinearity was not a significant factor among the

Figure 3 Map of Confirmed Cases of COVID-19 (per 100,000 population) in the Conterminous 48 States.



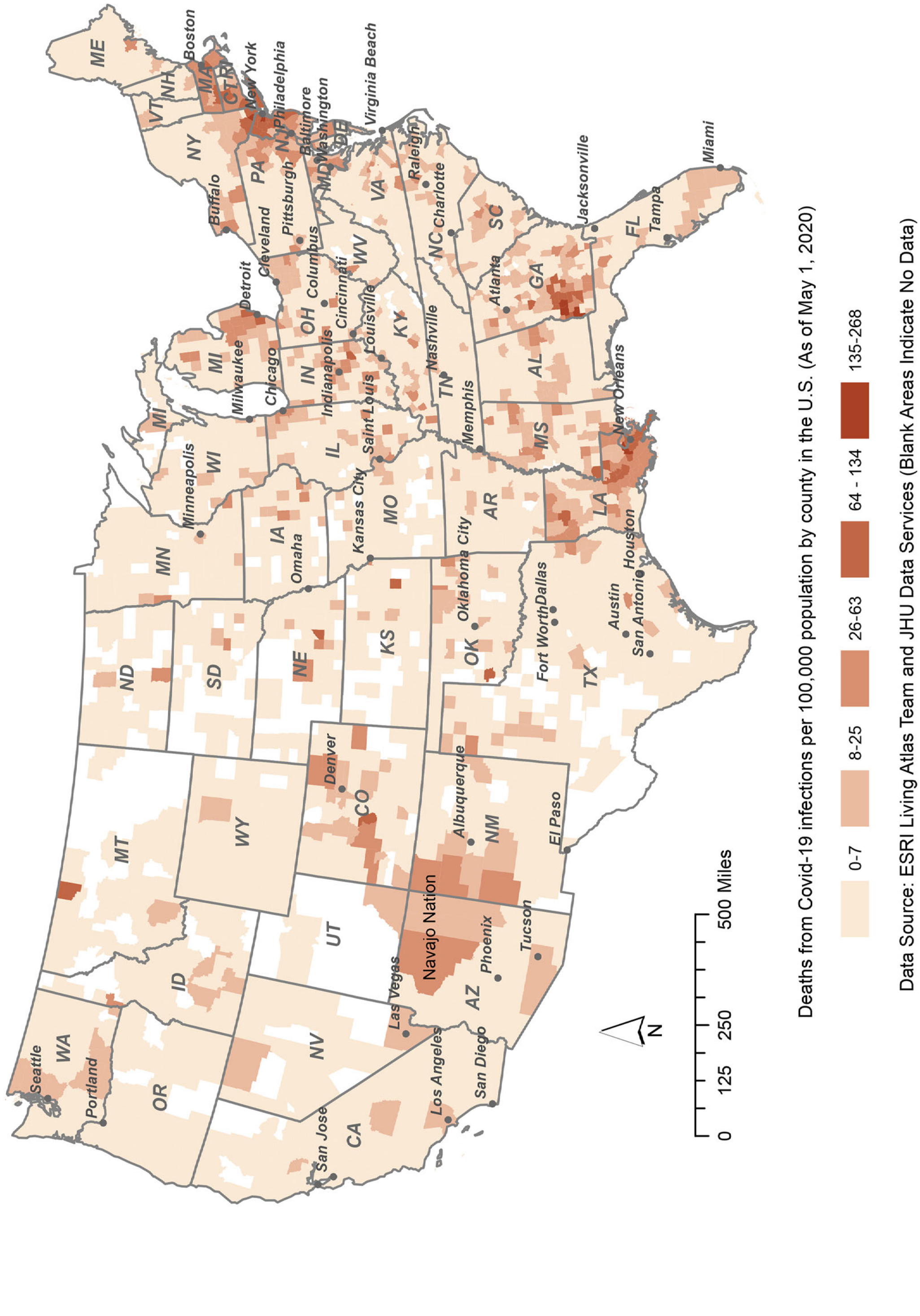


Figure 4 Map of Deaths from COVID-19 (per 100,000 population) in the Conterminous 48 States.

Table 3 Top 25 Counties with the Highest Infection Rates from COVID-19 (Rural Counties are Shaded)

Ranking	County, State	USDA Locality	Population	Infections Per 100,000
1	Lincoln, Arkansas	Small MSA	13,695	5,943.8
2	Bledsoe, Tennessee	Rural or nonmetro	14,602	4,067.9
3	Rockland, New York	Large MSA	323,686	3,617.1
4	Nobles, Minnesota	Rural or nonmetro	21,839	3,397.6
5	Marion, Ohio	Rural or nonmetro	65,344	3,360.7
6	Dakota, Nebraska	Small MSA	20,317	3,228.8
7	Cass, Indiana	Rural or nonmetro	38,084	3,056.4
8	Pickaway, Ohio	Large MSA	57,420	3,016.4
9	Westchester, New York	Large MSA	968,815	2,990.3
10	Nassau, New York	Large MSA	1,356,564	2,643.0
11	Passaic, New Jersey	Large MSA	504,041	2,469.8
12	Louisa, Iowa	Rural or nonmetro	11,223	2,396.9
13	Seward, Kansas	Rural or nonmetro	22,692	2,322.4
14	Orange, New York	Large MSA	378,227	2,287.0
15	Union, New Jersey	Large MSA	553,066	2,274.2
16	Suffolk, New York	Large MSA	1,487,901	2,262.5
17	Blaine, Idaho	Rural or nonmetro	21,994	2,259.7
18	Randolph, Georgia	Rural or nonmetro	7,087	2,243.5
19	Hudson, New Jersey	Large MSA	668,631	2,230.8
20	Terrell, Georgia	Small MSA	8,859	2,065.7
21	Early, Georgia	Rural or nonmetro	10,348	2,058.4
22	New York City, New York	Large MSA	8,443,713	1,983.5
23	Ford, Kansas	Rural or nonmetro	34,484	1,957.4
24	Dawson, Nebraska	Rural or nonmetro	23,204	1,915.6
25	St. John the Baptist, Louisiana	Large MSA	43,446	1,767.7

independent variables included in the model as the variance inflation factor (VIF) values for each independent variable were as low as <2 (data not shown).

Results of a sensitivity analysis found that when only rural counties were used, population density became insignificant ($b = 0.100$; Std error $= 0.072$) for infections and the model explained less than 1% of the variations of infections (Adjusted R square $= 0.001$). This finding indicates that population density is less likely to be a risk factor for rural counties because rural areas in general have much lower population densities in comparison to their metropolitan counterparts.

The results of regression model using COVID-19 deaths as the dependent variable are similar to those observed in the model for confirmed cases. Population density remained the strongest predictor of COVID-19 mortality rates. The variables, percent older population and poverty, both showed expected significant and positive associations with COVID-19 deaths. This is consistent with the hypothesis that the elderly and the poor are at greater risks of COVID-19 deaths. The percent of the minority population was significant but negative. One possible explanation is that the data on coronavirus deaths are not disaggregated by race or ethnicity and Whites were the majority in the death tolls for most counties. Percent of the uninsured population was positive but was not significant, which could be a result of underestimates or undercounting of COVID-19 induced fatalities, especially for those disadvantaged communities (ie, inner cities or remote rural communities). The regression model explains 39% of the variations of COVID-19 across 1,624 counties. Results from a sensitivity analysis using rural counties only and regressing mortality against population density alone were similar to the findings for infections.

Discussion and Conclusions

This article analyzes country-wide data from the ongoing COVID-19 pandemic with a focus on spatial disparities among metropolitan and nonmetropolitan communities. As the COVID-19 epidemic in most areas of the US continues to mount, these findings must be considered an interim appraisal. More definitive conclusions can only be reached once the epidemic has abated and when other relevant data, some not presently available, are included. However, several notable trends are apparent. Large, densely populated cities and their surrounding

Table 4 Top 25 Metropolitan Areas Ranked by Death Rate per 100,000 Population

Ranking	Metropolitan areas	MSA Size	Population	Death per 100,000
1	New York-Newark-Jersey City, NY-NJ-PA	Large	19,990,592	146.9
2	Albany, GA	Small	153,101	113.0
3	Bridgeport-Stamford-Norwalk, CT	Medium	944,348	85.8
4	New Orleans-Metairie, LA	Large	1,263,635	83.3
5	Detroit-Warren-Dearborn, MI	Large	4,317,179	72.8
6	Hartford-West Hartford-East Hartford, CT	Large	1,209,367	67.2
7	Springfield, MA	Medium	630,275	63.9
8	Trenton, NJ	Medium	368,762	62.6
9	New Haven-Milford, CT	Medium	859,339	59.6
10	Boston-Cambridge-Newton, MA-NH	Medium	4,811,732	54.5
11	Flint, MI	Medium	409,361	45.9
12	Grand Island, NE	Small	84,729	41.3
13	Houma-Thibodaux, LA	Small	210,801	36.5
14	Farmington, NM	Small	127,455	35.3
15	Baton Rouge, LA	Medium	829,642	32.8
16	East Stroudsburg, PA	Small	167,586	32.2
17	Greeley, CO	Medium	295,123	32.2
18	Lewiston, ID-WA	Small	62,492	32.0
19	Allentown-Bethlehem-Easton, PA-NJ	Medium	834,615	31.8
20	Lancaster, PA	Medium	538,347	31.6
21	Shreveport-Bossier City, LA	Medium	441,339	31.5
22	Scranton–Wilkes-Barre–Hazleton, PA	Medium	556,926	30.9
23	Saginaw, MI	Small	192,778	30.6
24	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	Large	6,069,448	30.5
25	Flagstaff, AZ	Small	140,217	30.0

metropolitan areas are the hotspots of the COVID-19 pandemic both in terms of incidence and mortality. In general, the larger metropolitan areas, including both urban and suburban areas, have higher incidence and mortality rates than smaller and nonmetropolitan areas. However, as of May 1, nonmetropolitan areas that span 1,667 counties in the United States have amassed nearly 6% (or 60,000) of confirmed cases and 4% (or over 2,300) of deaths from COVID-19. It is striking that the infection rates in some small metropolitan areas (eg, Albany, GA; Pine Bluff, AR; Grand Island, NE) and numerous rural counties (eg, Bledsoe, TN; Nobles, MN; and Marion, OH) ranked very high nationally. Likewise, the mortality rates in some small metropolitan areas (eg, Albany, GA; Grant Island, NE; Houma-Thibodaux, LA; and Farmington, NM) and nonmetropolitan counties (eg, Randolph, Early, and Mitchell counties in southwest Georgia and Toole County in Montana) actually exceed those of the larger cities. The complete explanation for this paradox is not apparent but likely involves chance or "local" factors that become important in smaller communities. For example, the high rates in Albany, Georgia, have been traced to the presence of an infected individual at 2 funerals that were highly attended.[39] The introduction of an infected individual thus can have a disproportionate effect on disease rates when the population is relatively small but densely aggregated. Additionally, it is likely that the limited health care facilities of small cities and rural communities are a critical bottleneck to access to services. Many rural areas tend to have higher poverty rates and much older populations with a higher prevalence of comorbidities. These populations may be more vulnerable to symptomatic and more serious courses of disease than other populations.[40]

Regression analyses support the hypotheses of correlations between COVID-19 infections and deaths and the selected socioeconomic contextual factors including population density, proportions of elderly residents, poverty, and testing at the state level. It is reasonable to suggest that densely populated areas and communities that are aging, poverty-stricken, and lack testing are more at risk of COVID-19. Moreover, results from the sensitivity analyses corroborate that population density is a more effective predictor of COVID-19 infections and mortality for metropolitan areas, not for rural areas. The unexpected results for other socioeconomic risk factors, including proportion of minority and uninsured population, are largely due to the use of disaggregated data, undertesting, or underreporting. Despite this weakness, our findings corroborate anecdotal observations and

Table 5 Top 25 Counties with the Highest Mortality Rates from COVID-19 (Rural Counties are Shaded)

Ranking	County, State	USDA Locality	Population	Deaths per 100,000
1	Randolph, Georgia	Rural or Nonmetro	7,087	268.10
2	Terrell, Georgia	Small MSA	8,859	214.47
3	New York City, New York	Large MSA	8,443,713	213.99
4	Early, Georgia	Rural or Nonmetro	10,348	193.27
5	St. John the Baptist, Louisiana	Large MSA	43,446	158.82
6	Rockland, New York	Large MSA	323,686	157.87
7	Essex, New Jersey	Large MSA	793,555	149.45
8	Mitchell, Georgia	Rural or Nonmetro	22,432	133.74
9	Dougherty, Georgia	Small MSA	91,049	132.90
10	Nassau, New York	Large MSA	1,356,564	125.32
11	Union, New Jersey	Large MSA	553,066	124.76
12	Bergen, New Jersey	Large MSA	929,999	122.15
13	Toole, Montana	Rural or Nonmetro	4,976	120.58
14	Hudson, New Jersey	Large MSA	668,631	119.35
15	Passaic, New Jersey	Large MSA	504,041	113.68
16	Orleans, Louisiana	Large MSA	389,648	111.38
17	Westchester, New York	Large MSA	968,815	106.42
18	Decatur, Indiana	Rural or Nonmetro	26,552	105.45
19	Wilcox, Georgia	Rural or Nonmetro	8,846	101.74
20	Wayne, Michigan	Large MSA	1,761,382	101.17
21	Greer, Oklahoma	Rural or Nonmetro	5,943	100.96
22	Turner, Georgia	Rural or Nonmetro	7,962	100.48
23	Clay, Georgia	Rural or Nonmetro	3,001	99.97
24	Coffey, Kansas	Rural or Nonmetro	8,296	96.43
25	Bienville, Louisiana	Rural or Nonmetro	13,668	95.11

Table 6 Results of Regression Analysis of Confirmed Cases and Deaths of COVID-19 ($n = 1{,}624$)

	Model for Incidence		Model for Deaths	
Independent Variables	Coefficient	Standard Error	Coefficient	Standard Error
Constant	-2940.18^{***}	530.334	-296.338^{***}	54.516
Population density	2.433^{***}	0.072	0.231^{***}	0.007
% Aged 65 years old and above	62.149^{*}	24.387	6.576^{**}	2.507
% Population below poverty	$-.472$	19.223	4.107^{*}	1.976
% Minority population	-2.702	6.067	-1.801^{**}	0.624
% Uninsured population	34.607	26.222	3.974	2.696
% Of population tested (statewide)	690.644^{***}	103.783	43.423^{***}	10.668
Adjusted R Square	0.447 ($F = 219.437$, $P < .001$)		0.391 ($F = 174.656$, $P < .001$)	

All statistical tests were 2-sided and a P value $\leq .05$ was considered significant.
*Significance level of .05.
**Significance level of .01.
***Significance level of .001. All variance inflation factor (VIF) values for the independent variables were below 2.0.

empirical research that highlight the vulnerability of those disadvantaged communities.[11,38] Our findings about the geographic heterogeneity and sociodemographic correlates of the COVID-19 outbreak can inform policy and government efforts to address the uneven distribution of COVID-19 cases across demographic groups and geographic areas.[38] Given the fact of implementations of nationwide lockdown and social distancing practices, the pandemic may be showing signs of receding in large metropolitan areas, but it is predicted to continue to diffuse from cities to small places and rural communities. These findings support the provision of additional testing facilities and services, including mobile health clinics and telehealth services, to help the rural communities that have been underserved and undertested during this pandemic.[41-43]

It is important to note that these data are subject to several methodological limitations, including the ecologic nature of the analysis and biases that are related to the under- (and sometimes over-) ascertainment of cases. First, the primary data in this spatial study are counties, not individuals. Thus, we cannot validly conclude, for example, that because counties with a high prevalence of population aged 65 or above have higher mortality rates that individuals who are 65 or older will do so. Furthermore, due to the unavailability of disaggregated data, our findings were based on analysis of total counts of infections and deaths in each county, prohibiting us from identifying disparities between demographic groups (eg, gender, age, or race/ethnicity) within the population.[44] It would obviously be valuable to conduct more in-depth analysis using individual-level data to examine racial/ethnic disparities in infection and mortality rates if data on racial/ethnic breakdowns are released by health agencies at later stages. Our focus in this early analysis is to identify regions with geographic disparities in order to probe the underlying causes behind these. Thus, the ecologic approach is a valid and appropriate first step in this process.

Ascertainment bias is another important, potential limitation. The incidence rate obviously depends on the frequency of testing. Shortages of testing equipment remain a critical problem for both disease control and epidemiological analysis. Therefore, low prevalence rates in many areas, especially sparsely populated rural counties, are likely artifacts of lower rates of testing, as validated by the positive association of infections and testing rate at the state level in our regression analysis.[38] Thus, underascertainment should disproportionately reflect poorer, more rural counties, whose true prevalence rates will likely be higher than recorded. Conversely, the data are also susceptible to inflation of mortality rates that can occur when individuals who are ill from etiologically unrelated diseases succumb to these diseases while they are infected. In large metropolitan areas, such as New York City, thousands of individuals may die each day due to noncoronavirus causes, such as lung cancer, cardiovascular disease, etc. In many individuals, the coronavirus infection doubtlessly exacerbated their pre-existing illness and, in this way, causally contributed to their deaths. However, deaths from these underlying causes in the presence of a positive COVID-19 serological test is likely to be recorded as a "coronavirus death" even if the contribution of the virus to death was modest. That is, some deaths that under nonpandemic circumstances would be attributable to noncoronavirus causes likely will be counted as deaths due to coronavirus. The actual autopsy rate in the present health crisis is expected to be exceptionally low. This type of ascertainment bias likely would have complex, and different effects in rural than in metropolitan areas. For example, in rural areas, a lower rate of testing would fail to identify some decedents as coronavirus cases and likely would lead to an underestimate of the effect of the virus on mortality.

In summary, our analysis found that COVID-19 infections and deaths are (not surprisingly highly) concentrated in large cities and their surrounding metropolitan areas. However, equivalently high rates were found in small cities and nonmetropolitan or rural communities. Communities with high proportions of elder populations, high poverty rates, and high population density are found at high risk of COVID-19. Moreover, the magnitude of infections and deaths is positively associated with testing rates at the state level, which supports expanding testing capacities to undertested areas, especially rural communities, to stem the spread of the virus. These findings are subject to ascertainment biases and should be considered provisional guides to the geography and underlying risk factors of the pandemic in the United States.

References

1. Mayo Clinic. Coronavirus disease 2019 (COVID-19)—Symptoms and causes 2020. Available at: https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963. Accessed April 14, 2020.
2. Holshue ML, DeBolt C, Lindquist S, et al. First case of 2019 novel coronavirus in the United States. *N Engl J Med*. 2020;382:929-936.
3. Lewis D. Is the coronavirus airborne? Experts can't agree. *Nature*. 2020;580:175.
4. Bourouiba L. Turbulent gas clouds and respiratory pathogen emissions: potential implications for reducing transmission of COVID-19. *JAMA*. 2020;323(18):1837-1838.
5. Gautret P, Lagier J-C, Parola P, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *Int J Antimicrob Agents*. 2020;105949. https//doi.org/10.1016/j.ijantimicag.2020.105949
6. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*. 2020;323(13):1239-1242.
7. Peeri NC, Shrestha N, Rahman MS, et al. The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? *Int J Epidemiol*. 2020;dyaa033. https//doi.org/10.1093/ije/dyaa033

8. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun*. 2020;109:102433.
9. Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol*. 2020;92(4):441-447.
10. Lighter J, Phillips M, Hochman S, et al. Obesity in patients younger than 60 years is a risk factor for Covid-19 hospital admission. *Clin Infect Dis*. 2020;ciaa415. https://doi.org/10.1093/cid/ciaa415
11. Wadhera RK, Wadhera P, Gaba P, et al. Variation in COVID-19 hospitalizations and deaths across New York City boroughs. *JAMA*. 2020. https//doi.org/10.1001/jama.2020.7197
12. Newkirk VR II. The coronavirus's unique threat to the south: more young people in the south seem to be dying from covid-19. Why? *Atlantic*. 2020. Available at: https://www.theatlantic.com/politics/archive/2020/04/coronavirus-unique-threat-southyoung-people/609241/
13. Krout JA. Rural ageing. *GIA— Grantmakers in aging*. Available at https://www.giaging.org/. Accessed April 8, 2020.
14. Baernholdt M, Yan G, Hinton I, Rose K, Mattos M. Quality of life in rural and urban adults 65 years and older: findings from the National Health and Nutrition Examination survey. *J Rural Health*. 2012;28(4):339-347.
15. Lakhani A. Introducing the percent, number, availability, and capacity [PNAC] spatial approach to identify priority rural areas requiring targeted health support in light of COVID-19: a commentary and application. *J Rural Health*. 2020. https//doi.org/10.1111/jrh.12436
16. Befort CA, Nazir N, Perri MG. Prevalence of obesity among adults from rural and urban areas of the United States: findings from NHANES (2005–2008). *J Rural Health*. 2012;28(4):392-397.
17. Monnat S. Why coronavirus could hit rural areas harder. *Lerner Center for Public Health Promotion, Syracuse University*; 2020. Available at: https://lernercenter.syr.edu/2020/03/24/why-coronavirus-could-hit-rural-areas-harder/. Accessed April 9, 2020.
18. Ranscombe P. Rural areas at risk during COVID-19 pandemic. *Lancet Infect Dis*. 2020;20:545.
19. Kamel Boulos MN, Geraghty EM. Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. *Int J Health Geogr*. 2020;19(1):8.
20. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis*. 2020;20(5):533-534.
21. Zhou C, Su F, Pei T, et al. COVID-19: challenges to GIS with big data. *Geogr Sustain*. 2020;1(1):77-87.
22. Xu B, Gutierrez B, Mekaru S, et al. Epidemiological data from the COVID-19 outbreak, real-time case information. *Sci Data*. 2020;7(1):106.
23. Lipsitch M, Swerdlow DL, Finelli L. Defining the epidemiology of Covid-19—studies needed. *N Engl J Med*. 2020;382:1194-1196.
24. CDC COVID-19 Response Team. Geographic differences in covid-19 cases, deaths, and incidence—United States, February 12–April 7, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69:465-471.
25. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospitalization and critical illness among 4,103 patients with COVID-19 disease in New York City. *medRxiv*. https//doi.org/10.1101/2020.04.08.20057794
26. Chen R, Liang W, Jiang M, et al. Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in China. *Chest*. 2020. https//doi.org/10.1016/j.chest.2020.04.010
27. Wu C, Chen X, Cai Y, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with Coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med*. 2020. https//doi.org/10.1001/jamainternmed.2020.0994
28. Martin SL, Kirkner GJ, Mayo K, et al. Urban, rural, and regional variations in physical activity. *J Rural Health*. 2005;21(3):239-244.
29. Moss JL, Liu B, Feuer EJ. Urban/rural differences in breast and cervical cancer incidence: the mediating roles of socioeconomic status and provider density. *Women's Health Issues*. 2017;27(6):683-691.
30. Lawrence E, Hummer RA, Harris KM. The cardiovascular health of young adults: disparities along the urban-rural continuum. *Ann Am Acad Pol Soc Sci*. 2017;672(1):257-281.
31. Srinivasan S, Moser RP, Willis G, et al. Small is essential: importance of subpopulation research in cancer control. *Am J Public Health*. 2015;105(Suppl 3):S371-S373.
32. Florida R. The geography of coronavirus. *Citylab*. 2020. Available at: https://www.citylab.com/equity/2020/04/coronavirus-spread-map-city-urban-density-suburbs-rural-data/609394/. Accessed April 9, 2020.
33. Emanuel EJ, Persad G, Upshur R, et al. Fair allocation of scarce medical resources in the time of Covid-19. *N Engl J Med*. 2020;382:2049-2055.
34. Benz JK, Espinosa O, Welsh V, Fontes A. Awareness of racial and ethnic health disparities has improved only modestly over a decade. *Health Aff*. 2011;30(10):1860-1867.
35. James CV, Moonesinghe R, Wilson-Frederick SM, Hall JE, Penman-Aguilar A, Bouye K. Racial/ethnic health disparities among rural adults—United States, 2012–2015. *MMWR Surv Summaries*. 2017;66(23):1-9.
36. Jackson CS, Oman M, Patel AM, Vega KJ. Health disparities in colorectal cancer among racial and ethnic minorities in the United States. *J Gastrointest Oncol*. 2016;7(Suppl 1):S32-S43.

37. National Academies of Sciences E, Medicine. *Communities in Action: Pathways to Health Equity*. Washington, D.C.: National Academies Press; 2017.
38. Souch JM, Cossman JS. A commentary on rural-urban disparities in covid-19 testing rates per 100,000 and risk factors. *J Rural Health*. 2020. https//doi.org/10.1111/jrh.12450
39. Willis H, Williams V. A funeral is thought to have sparked a covid-19 outbreak in Albany, Ga.—and led to many more funerals. *The Washington Post*, 2020. Available at: https://www.washingtonpost.com/politics/a-funeral-sparked-a-covid-19-outbreak-and-led-to-many-more-funerals/2020/04/03/546fa0cc-74e6-11ea-87da-77a8136c1a6d_story.html. Accessed 12 April 2020.
40. McLaughlin DK, Jensen L. Poverty among older Americans: the plight of nonmetropolitan elders. *J Gerontol*. 1993;48(2):S44-S54.
41. Nagata JM. Rapid scale-up of Telehealth during the COVID-19 pandemic and implications for subspecialty care in rural areas. *J Rural Health*. 2020. https//doi.org/10.1111/jrh.12433
42. Prusaczyk B. Strategies for disseminating and implementing COVID-19 public health prevention practices in rural areas. *J Rural Health*. 2020. https//doi.org/10.1111/jrh.12432
43. Adunlin G, Murphy PZ, Manis M. COVID-19: how can rural community pharmacies respond to the outbreak? *J Rural Health*. 2020. https//doi.org/10.1111/jrh.12439
44. Laurencin CT, McClinton A. The covid-19 pandemic: a call to action to identify and address racial and ethnic disparities. *J Racial Ethn Health Disparities*. 2020;7(3):398-402.