Spatiotemporal Characteristics of the COVID-19 Epidemic in the United States

Yun Wang,[1] Ying Liu,[2,3,a] James Struthers,[1] and Min Lian[1,3]

[1]Division of General Medical Sciences, Department of Medicine, Washington University School of Medicine, St Louis, Missouri, USA, [2]Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St Louis, Missouri, USA, and [3]Alvin J. Siteman Cancer Center, Barnes-Jewish Hospital and Washington University School of Medicine, St Louis, Missouri, USA

Background. A range of near-real-time online/mobile mapping dashboards and applications have been used to track the coronavirus disease 2019 (COVID-19) pandemic worldwide; however, small area-based spatiotemporal patterns of COVID-19 in the United States remain unknown.

Methods. We obtained county-based counts of COVID-19 cases confirmed in the United States from 22 January to 13 May 2020 (N = 1 386 050). We characterized the dynamics of the COVID-19 epidemic through detecting weekly hotspots of newly confirmed cases using Spatial and Space-Time Scan Statistics and quantifying the trends of incidence of COVID-19 by county characteristics using the Joinpoint analysis.

Results. Along with the national plateau reached in early April, COVID-19 incidence significantly decreased in the Northeast (estimated weekly percentage change [EWPC]: −16.6%) but continued increasing in the Midwest, South, and West (EWPCs: 13.2%, 5.6%, and 5.7%, respectively). Higher risks of clustering and incidence of COVID-19 were consistently observed in metropolitan versus rural counties, counties closest to core airports, the most populous counties, and counties with the highest proportion of racial/ethnic minorities. However, geographic differences in incidence have shrunk since early April, driven by a significant decrease in the incidence in these counties (EWPC range: −2.0%, −4.2%) and a consistent increase in other areas (EWPC range: 1.5–20.3%).

Conclusions. To substantially decrease the nationwide incidence of COVID-19, strict social-distancing measures should be continuously implemented, especially in geographic areas with increasing risks, including rural areas. Spatiotemporal characteristics and trends of COVID-19 should be considered in decision making on the timeline of re-opening for states and localities.

Keywords. COVID-19; epidemiology; geography; clustering; spatiotemporal trend.

Since the first cluster of the coronavirus disease 2019 (COVID-19) was reported [1, 2], the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has triggered massive outbreaks and then evolved to a worldwide pandemic of COVID-19. As of 13 May 2020, 4 347 018 confirmed cases and 297 197 COVID-19–related deaths have been reported worldwide [3]. In the United States, the first COVID-19 case was reported on 21 January 2020 [4], and the national outbreak of COVID-19 beginning in early March of 2020 has caused 1 386 050 confirmed cases and 83 167 deaths from COVID-19 as of 13 May [5]. It is urgent to "flatten the epidemic curve" for COVID-19 in the United States.

Remarkable efforts have been made to map the coronavirus spread using near-real-time interactive online/mobile geographic information systems (GIS) dashboards, websites, and applications in and out of the United States [3, 5–7]. These maps provide timely information on descriptive statistics of the outbreak situation. However, no studies have comprehensively assessed small area-based characteristics of the spread of COVID-19 in the United States. Using government record–based surveillance data, we examined the spatiotemporal variations in COVID-19 as well as its associated geographic characteristics across the country. The results would enhance our understanding of small area-based spatiotemporal dynamics of the COVID-19 outbreak, and thus help inform multilevel strategies to control the spread of coronavirus and appropriate allocations of public health and healthcare resources in the United States.

METHODS

Data Source

We obtained the counts of COVID-19 cases diagnosed from 22 January to 13 May 2020 in the United States from the USAFacts, a not-for-profit initiative standardizing and providing the

publicly available government record–based data [5]. The daily-updated numbers were cumulated to form a time-series database of confirmed COVID-19 cases across all the US counties. The study is exempted from the ethics review due to the use of a publicly accessible data source.

County-level Variables

To identify the characteristics of counties with a high burden of COVID-19, we examined county-level geographic and sociodemographic factors, including rural-urban context, distance to the nearest core airport, population density, percentage of racial/ethnic minority population, percentage of the population 65 years or older, and percentage of the population below the federal poverty line. Using the Rural-Urban Continuum Codes of US Department of Agriculture [8], rural-urban context was defined as metropolitan (codes 1–3), urban (codes 4–7), and rural (codes 8–9) areas. There are 30 core airports with the highest volume of traffic across the country [9]. The Euclidean distance from the population-weighted centroid of a given county to its nearest core airport was calculated to measure the spatial relationship of that county with core airports. Population density was computed as the population number per square miles of land. County-level information on land areas, population sizes, and 3 other socioeconomic variables was retrieved from the combined 2014–2018 American Community Surveys to reduce the potential marginal error of the survey.

Statistical Analysis

We first created an epidemic curve to visualize the progression of newly confirmed COVID-19 cases by 4 US government–defined regions (Northeast, Midwest, South, and West) over 11 distinct time periods from 22 January through 13 May 2020, including the first 6 epi-weeks in combination (22 January–4 March) and individual epi-weeks from 5 March to 13 May.

Using Spatial and Space-Time Scan Statistics (SaTScan) [10, 11], we examined spatiotemporal clustering of confirmed COVID-19 cases through detection of the higher-than-expected geographic hotspots across the country. The SaTScan applies a predefined circular window with varied sizes and time periods to scan the study area and identify the most likely clusters of the event of interest using a space-time permutation statistical model, and uses a Monte Carlo simulation approach to generate 999 random datasets in computing the statistic for the statistical inference of a cluster. In this study, we defined the parameters of the scanning window as 150 miles of maximum geographic radius and the day as the minimum temporal scanning unit. Geographic clustering was detected in each of 11 time periods to characterize the dynamics of geographic hotspots of newly confirmed COVID-19 cases. The most likely high-risk clusters/hotspots were captured based on the Monte Carlo rank with $P < .05$. We further examined the associations of county

characteristics with COVID-19 clustering using logistic regressions. The outcome was whether or not a given county was identified as part of a hotspot. The analysis was performed separately for each of the 7th–16th epi-weeks. Considering the collinearity between county characteristics, county-level variables were not mutually adjusted for. Statistical significance was tested as 2-sided with $P < .05$.

Finally, we computed the incidence rates of COVID-19 during each of 11 time periods to quantify the overall risk of infection and spatiotemporal trend of the spread of COVID-19 by region and geographic/demographic characteristics. Temporal trends in COVID-19 incidence, overall and by region and county characteristics, were quantified by the estimated weekly percentage changes (EWPCs) using the Joinpoint regression [12].

The management of the database and logistic regressions were performed using the SAS System (version 9.4; SAS Institute Inc, Cary, NC). The Space-Time cluster analyses were performed using the SaTScan software (version 9.6; Martin Kulldorff, Harvard Medical School, Boston and Information Management Services Inc, Calverton, MD). The detected geographic hotspots were mapped using the ArcGIS package (version 10.2.2; ESRI, Redlands, CA). The temporal trend analyses were performed using the Joinpoint package (version 4.8.0.1; Statistical Methodology and Application Branch, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute).

RESULTS

Temporal Trend and Regional Differences of Confirmed Cases

As of 13 May 2020, a total of 1 386 050 COVID-19 cases were confirmed in the United States over 16 epi-weeks. Figure 1A shows the overall temporal trend of weekly counts of newly confirmed COVID-19 cases by 4 US regions. COVID-19 had occurred sporadically until early March (first 6 epi-weeks); 116 confirmed cases were reported mainly in the West region. The number of weekly confirmed COVID-19 cases subexponentially increased across the country from the 7th to the 11th epi-week, and slowly decreased in the following 5 epi-weeks. During the entire observation period, the largest proportion of cases was from the Northeast (48.6%), followed by the South (22.3%), Midwest (18.3%) and West (10.8%). During epi-weeks 11–16, the proportion of confirmed cases decreased from 55.3% to 29.1% in the Northeast but increased in 3 other regions (Midwest: 14.1% to 26.4%; South: 21.5% to 30.8%; and West: 9.0% to 13.7%).

Spatiotemporal Clustering and the Associated County-level Characteristics

Figure 2 shows the clustering dynamics of COVID-19. In the first 6 epi-weeks, 2 geographic clusters (covering 20 counties) were detected in southern and northern California. In the seventh epi-week, 6 geographic hotspots (covering 200 counties) were identified in Washington, New York, Massachusetts, and
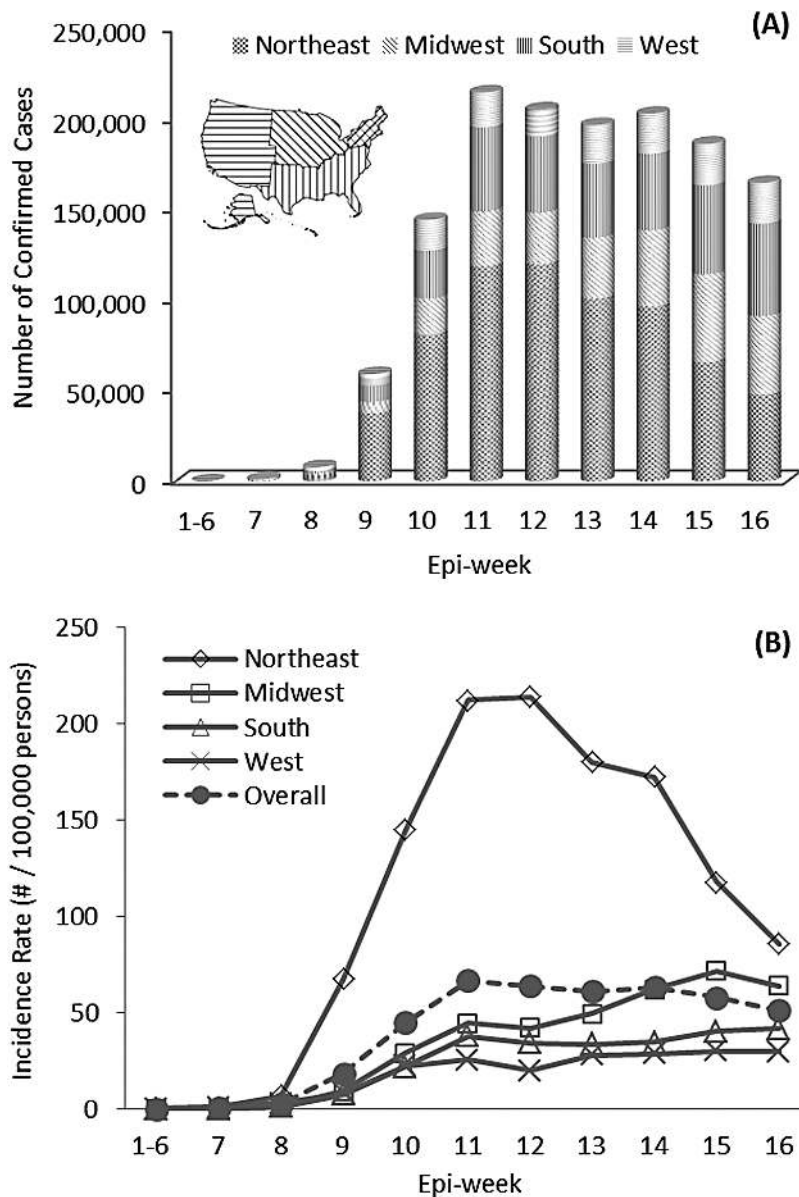
Figure 1. The temporal trend of weekly number of confirmed cases (*A*) and weekly incidence (*B*) of COVID-19 across 4 geographic regions in the United States over 16 epi-weeks, 22 January–13 May 2020. Abbreviation: COVID-19, coronavirus disease 2019.

northern Florida. In the following epi-weeks, many more geographic hotspots were detected across the country, including 19 in 783 counties, 41 in 1214 counties, 49 in 1048 counties, 47 in 1485 counties, 63 in 1355 counties, 66 in 1252 counties, 55 in 1356 counties, 71 in 1687 counties, and 66 in 1613 counties from the 8th to 16th epi-weeks, respectively. The results from geographic clustering analysis are robust. Starting in the eighth epi-week, the top-20 clusters were highly significant ($P < 0.001$) and came largely from the Midwest and South regions.

Table 1 shows the associations of county characteristics with COVID-19 clustering in the 7th to 16th epi-weeks. Compared with rural counties, the odds ratio (OR) of COVID-19 clustering was 5.31 (95% confidence interval [CI], 2.90–9.72) in

metropolitan counties and 3.26 (95% CI, 1.76–6.03) in urban counties in the seventh epi-week, and the association gradually decreased through the 13th epi-week (OR, 1.04; 95% CI, .86–1.27 in metropolitan counties; OR, .88; 95% CI, .72–1.07 in urban counties) and slightly increased thereafter (epi-week 16: OR, 1.77; 95% CI, 1.45–2.15 in metropolitan counties; OR, 1.08; 95% CI, .90–1.31 in urban counties). The distance to the nearest core airports was strongly and inversely associated with COVID-19 clustering (epi-week 7: OR, 14.4; 95% CI, 5.75–35.8; epi-week 16: OR, 1.95; 95% CI, 1.59–2.38 for the lowest vs the highest quartiles). The population density was significantly and consistently associated with COVID-19 clustering from the seventh (OR, 6.69; 95% CI, 3.77–11.9
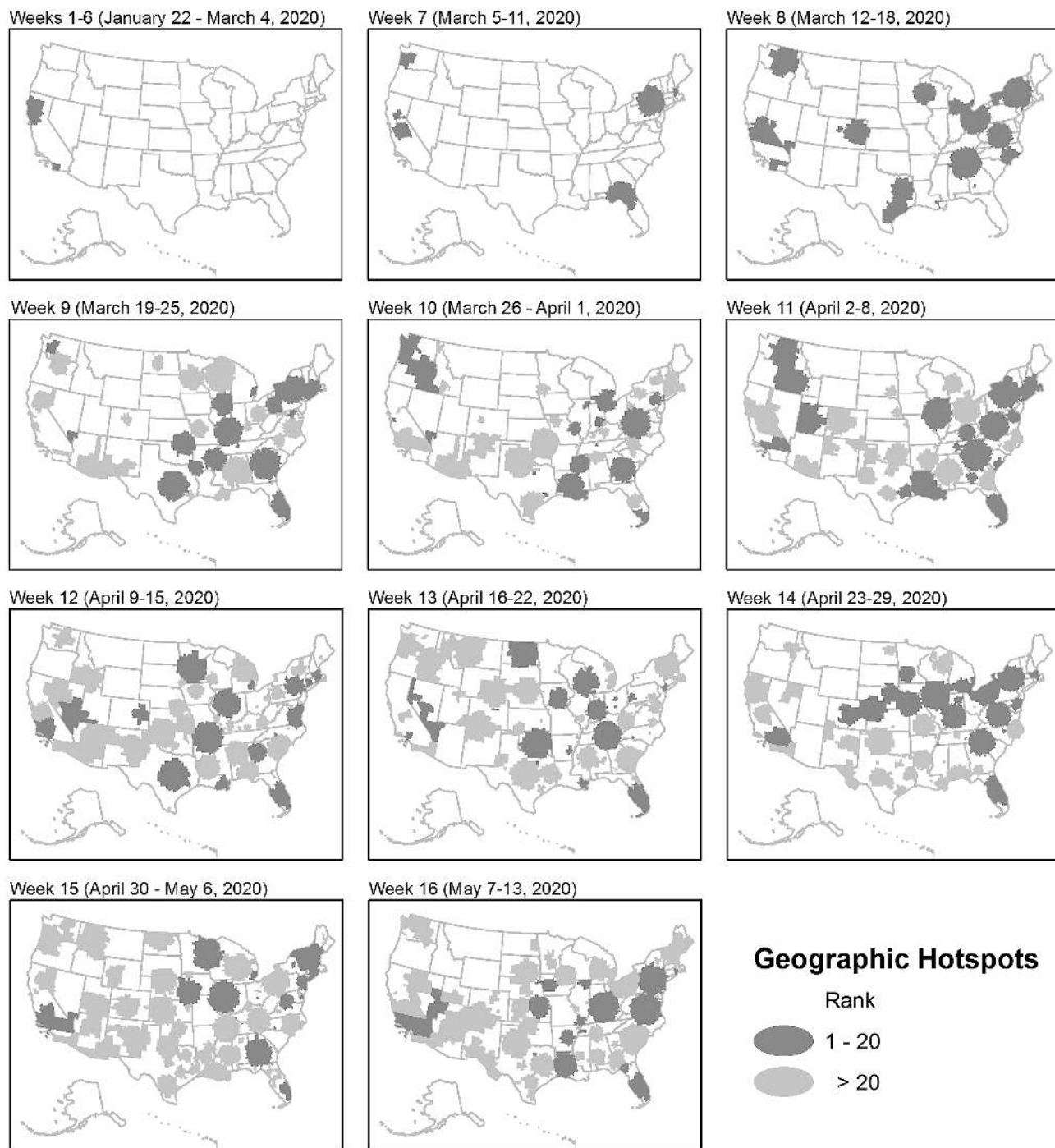
Figure 2. Weekly dynamics of geographic clustering of newly confirmed COVID-19 cases in the United States over 16 epi-weeks, 22 January–13 May 2020. Abbreviation: COVID-19, coronavirus disease 2019.

for highest vs lowest quartiles) through the 16th epi-weeks (OR, 2.56; 95% CI, 2.09–3.15). A significantly increased likelihood was also observed in areas with the highest versus lowest percentage of minorities from the seventh (OR, 6.93; 95% CI, 3.83–12.6) to the 16th (OR, 2.34; 95% CI, 1.91–2.87) epi-weeks, except for epi-weeks 13 and 14. A higher percentage of older people was associated with a lower likelihood of COVID-19 clustering from the eighth (OR, .57; 95% CI, .45–.72 for the highest vs lowest quartiles) to the 12th (OR, .63; 95% CI, .52–.77) epi-weeks and in the 16th epi-week (OR, .68; 95% CI, .56–.83). There was no clear pattern in the association between poverty and COVID-19 clustering.

Table 1. County-level Associations Between Geographic and Sociodemographic Characteristics and Risk of Confirmed COVID-19 Clustering in the United States From 5 March to 13 May 2020

Variable	Week 7 (5–11 March)	Week 8 (12–18 March)	Week 9 (19–25 March)	Week 10 (26 March—1 April)	Week 11 (2–8 April)	Week 12 (9–15 April)	Week 13 (16–22 April)	Week 14 (23–29 April)	Week 15 (30 April–6 May)	Week 16 (7–13 May)
Area										
Metro	5.31 (2.90–9.72)	3.63 (2.79–4.71)	3.40 (2.73–4.24)	3.16 (2.51–3.98)	3.05 (2.49–3.74)	1.93 (1.58–2.36)	1.04 (.86–1.27)	1.75 (1.43–2.13)	1.34 (1.10–1.63)	1.77 (1.45–2.15)
Urban	3.26 (1.76–6.03)	1.95 (1.49–2.54)	2.17 (1.74–2.70)	2.10 (1.67–2.64)	1.65 (1.35–2.01)	1.57 (1.29–1.92)	.88 (.72–1.07)	1.05 (.87–1.28)	1.01 (.83–1.22)	1.08 (.90–1.31)
Rural	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Distance to nearest core airports (quartile)										
Q1 (shortest)	14.4 (5.75–35.8)	24.6 (15.6–38.8)	4.63 (3.70–5.79)	2.30 (1.86–2.84)	4.67 (3.76–5.79)	2.75 (2.24–3.39)	1.03 (.84–1.26)	1.28 (1.04–1.56)	1.57 (1.28–1.91)	1.95 (1.59–2.38)
Q2	16.0 (6.43–39.8)	20.0 (12.7–31.7)	3.10 (2.48–3.88)	1.39 (1.12–1.73)	3.33 (2.69–4.13)	1.98 (1.61–2.44)	1.20 (.98–1.47)	1.18 (.96–1.44)	1.41 (1.15–1.72)	1.93 (1.58–2.36)
Q3	12.0 (4.79–30.2)	10.0 (6.30–16.0)	2.16 (1.72–2.72)	1.28 (1.02–1.59)	2.83 (2.28–3.50)	1.63 (1.32–2.00)	1.10 (.89–1.34)	1.21 (.99–1.48)	1.36 (1.12–1.66)	1.51 (1.23–1.84)
Q4 (longest)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Population density[a] (quartile)										
Q1 (lowest)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	3.27 (1.78–6.02)	3.20 (2.29–4.47)	4.23 (3.32–5.39)	1.93 (1.53–1.42)	2.38 (1.92–2.95)	1.93 (1.58–2.37)	1.06 (.87–1.30)	1.28 (1.04–1.57)	1.34 (1.10–1.64)	1.35 (1.10–1.65)
Q3	4.31 (2.38–7.80)	7.41 (5.39–10.2)	4.74 (3.72–6.03)	2.30 (1.83–2.89)	3.59 (2.90–4.45)	1.58 (1.28–1.94)	1.15 (.94–1.41)	1.39 (1.13–1.70)	1.54 (1.26–1.88)	1.62 (1.32–1.98)
Q4 (highest)	6.69 (3.77–11.9)	9.56 (6.97–13.1)	6.05 (4.75–7.70)	3.15 (2.51–3.94)	5.42 (4.36–6.75)	1.83 (1.49–2.24)	1.22 (1.00–1.49)	2.25 (1.84–2.76)	1.47 (1.20–1.79)	2.56 (2.09–3.15)
% Minority population[b] (quartile)										
Q1 (lowest)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	3.11 (1.64–5.87)	1.73 (1.35–2.22)	1.35 (1.08–1.70)	1.76 (1.38–2.23)	1.22 (.99–1.50)	1.28 (1.04–1.57)	.96 (.79–1.18)	.85 (.70–1.04)	1.03 (.84–1.25)	.99 (.81–1.21)
Q3	5.45 (2.98–9.96)	2.35 (1.84–2.99)	2.40 (1.93–2.98)	3.11 (2.47–3.92)	2.29 (1.86–2.81)	1.39 (1.13–1.71)	1.09 (.89–1.34)	1.23 (1.01–1.50)	1.23 (1.01–1.50)	1.43 (1.17–1.75)
Q4 (highest)	6.93 (3.83–12.6)	2.03 (1.59–2.60)	4.77 (3.83–5.94)	4.02 (3.20–5.06)	3.98 (3.22–4.91)	2.39 (1.95–2.93)	1.01 (.82–1.23)	1.01 (.83–1.24)	1.56 (1.28–1.91)	2.34 (1.91–2.87)
% Older population[c] (quartile)										
Q1 (lowest)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	1.27 (.87–1.86)	.91 (.73–1.13)	1.04 (.85–1.27)	1.07 (.88–1.32)	.79 (.65–.96)	.74 (.60–.90)	1.08 (.88–1.32)	1.04 (.85–1.27)	.97 (.80–1.19)	.88 (.72–1.07)
Q3	.92 (.61–1.38)	.91 (.73–1.13)	.91 (.75–1.12)	.74 (.60–.91)	.76 (.63–.93)	.90 (.74–1.10)	.93 (.76–1.14)	1.18 (.97–1.44)	.95 (.78–1.16)	.76 (.62–.93)
Q4 (highest)	.66 (.42–1.02)	.57 (.45–.72)	.63 (.51–.77)	.52 (.42–.65)	.46 (.37–.56)	.63 (.52–.77)	.99 (.81–1.21)	1.19 (.97–1.45)	.85 (.70–1.04)	.68 (.56–.83)
% Poverty[d] (quartile)										
Q1 (lowest)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Q2	1.70 (1.12–2.56)	1.12 (.90–1.41)	1.28 (1.04–1.57)	1.04 (.84–1.30)	1.00 (.82–1.22)	1.26 (1.03–1.55)	.89 (.72–1.09)	1.08 (.88–1.32)	1.06 (.87–1.30)	.84 (.69–1.02)
Q3	.81 (.50–1.31)	1.09 (.87–1.36)	1.26 (1.03–1.55)	1.43 (1.15–1.77)	1.33 (1.09–1.63)	1.07 (.87–1.31)	.94 (.76–1.15)	.82 (.67–1.01)	.96 (.79–1.17)	.77 (.63–.94)
Q4 (highest)	1.73 (1.15–2.61)	.68 (.54–0.87)	1.32 (1.07–1.62)	1.66 (1.35–2.05)	1.49 (1.22–1.82)	1.32 (1.08–1.62)	1.05 (.86–1.29)	.59 (.48–.73)	1.33 (1.09–1.63)	.89 (.73–1.09)

Data are presented as odds ratios (95% confidence interval). The regression analysis was performed for counties in 48 contiguous states and Washington, DC.

Abbreviations: COVID-19, coronavirus disease 2019; Q, quartile.

[a]The number of persons per square mile of lands.

[b]Percentage of racial and ethnic minorities.

[c]Percentage of persons aged 65 years or older.

[d]Percentage of persons below the federal poverty line.

Geographic Variation and Trend in Incidence

In the first 6 epi-weeks, COVID-19 cases were reported in 27 counties from the West Coast and Northeast states with the highest county-level incidence of 3.4 per 100 000 persons. Starting in the 7th through the 16th epi-weeks, SARS-CoV-2 spread to broad geographic areas. As of 13 May 92.0% of US counties had confirmed COVID-19 cases, and the median county-level cumulative incidence rate was 88.0 per 100 000 persons (interquartile range, 36.1–219.3/100 000 persons), with the highest reaching 14 426/100 000 persons (Supplementary Figure 1). The incidence of COVID-19 reached a peak in the Northeast in the 12th epi-week (214.2/100 000 persons), followed by a significant reduction of 16.6% weekly until the 16th epi-week. However, the incidence consistently increased in the Midwest, South, and West regions from the 10th to 16th epi-weeks, with significant EWPCs of 13.2%, 5.6%, and 5.7%, respectively. Overall, COVID-19 incidence reached the national plateau in epi-week 11 (66.6/10 000 persons), followed by a slight and nonsignificant decrease in the recent 5 weeks (Figure 1B, Table 2).

Figure 3 illustrates the trends in the incidence of COVID-19 by county characteristics. Over 16 epi-weeks, the incidence was significantly higher in metropolitan versus urban/rural areas, in areas closest to versus farthest from core airports, in the most versus least populous areas, and in areas with the highest versus lowest percentage of minorities and in those with the lowest versus highest percentage of the population aged 65 years and older (Figure 3A–E). The incidence dramatically increased from the seventh epi-week and reached a peak in the 11th epi-week in metropolitan areas (75.7/100 000 persons), counties closest to core airports (91.7/100 000 persons), most populous counties (79.6/100 000 persons), and counties with the highest percentage of minorities (100.9/100 000 persons), followed by a significant decrease thereafter (EWPCs = 2.0%, 2.8%, 2.6%, and 4.2%, respectively) (Table 2). Notably, the incidence consistently increased from epi-week 7 to 16 in rural (0.04 to 35.0/100 000 persons) and urban (0.1 to 37.5/100 000 persons) areas (Figure 3A). Unlike in metropolitan areas, the incidence continued to increase in rural and urban areas after the 11th epi-week, with a significant EWPC of 17.8% and 18.1%, respectively (Table 2). Similarly, a consistent increase in the incidence of COVID-19 from epi-week 7 to 16 was also observed in counties farther from core airports and in less populous counties and those with fewer minorities (Figure 3B–D, and Table 2). The incidence in counties with the lowest or highest percentage of elderly persons increased from epi-week 7 to 11 and remained steady thereafter. Overall, geographic disparities in the incidence of COVID-19 by county characteristics decreased from the 11th epi-week. There was no significant difference in the incidence of COVID-19 for the highest versus lowest percentage of population below the federal poverty line (Figure 3F).

Table 2. Temporal Trends of the Incidence of COVID-19 by Region and County Characteristics in the United States from 5 March to 13 May 2020

Variable	Trend 1		Trend 2	
	Epi-weeks	EWPC	Epi-weeks	EWPC
Overall	7–10	342.0*	10–16	−0.9
Region				
Northeast	7–11	111.4*	11–16	−16.6*
Midwest	7–10	385.2*	10–16	13.2*
South	7–10	399.8*	10–16	5.6*
West	7–10	184.0*	10–16	5.7*
Area				
Metro	7–10	353.6*	10–16	−2.0*
Urban	7–10	396.7	10–16	18.1*
Rural	7–11	167.9	11–16	17.8*
Distance to nearest core airports (quartile)				
Q1 (shortest)	7–10	352.4*	10–16	−2.8*
Q2	7–10	307.4*	10–16	8.4*
Q3	7–10	393.0*	10–16	12.9*
Q4 (longest)	7–10	420.8	10–16	1.5
Population density (quartile)				
Q1 (lowest)	7–10	275.8*	10–16	18.5*
Q2	7–10	344.8	10–16	20.3*
Q3	7–10	426.5*	10–16	16.9*
Q4 (highest)	7–10	354.5*	10–16	−2.6*
% Minority population (quartile)				
Q1 (lowest)	7–10	415.0*	10–16	7.2*
Q2	7–10	332.0*	10–16	7.8*
Q3	7–10	359.6*	10–16	4.2*
Q4 (highest)	7–10	336.6*	10–16	−4.2*
% Older population (quartile)				
Q1 (lowest)	7–10	332.3*	10–16	1.7
Q2	7–10	380.2*	10–16	−7.0*
Q3	7–10	529.0*	10–16	6.0*
Q4 (highest)	7–10	385.0*	10–16	1.7
% Poverty (quartile)				
Q1 (lowest)	7–10	347.9*	10–16	−2.3*
Q2	7–11	79.0	11–16	−4.0
Q3	7–10	400.4*	10–16	4.1*
Q4 (highest)	7–11	88.9*	11–16	−11.2

Epi-weeks 7 (5–11 March), 8 (12–18 March), 9 (19–25 March), 10 (26 March–1 April), 11 (2–8 April), 12 (9–15 April), 13 (16–22 April), 14 (23–29 April), 15 (30 April–6 May), 16 (6–13 May).
*$P < .05$ from the Joinpoint regressions.

Abbreviations: COVID-19, coronavirus disease 2019; EWPC, estimated weekly percentage changes from the Joinpoint regressions; Q, quartile.

DISCUSSION

Using a national time-series database of confirmed COVID-19 cases, we examined the spatiotemporal patterns of COVID-19 in the United States during the first 16 epi-weeks. COVID-19 cases sporadically occurred in the West Coast and Northeast states in the first 6 epi-weeks and increased rapidly across the country thereafter until the 11th epi-week, and then slightly decreased from the 12th epi-week. Despite a remarkable reduction in newly confirmed cases from the Northeast in the recent 4 weeks, the risk of SARS-CoV-2 infection continued to consistently increase in the Midwest, South, and West regions. Geographic clustering of COVID-19 was first identified
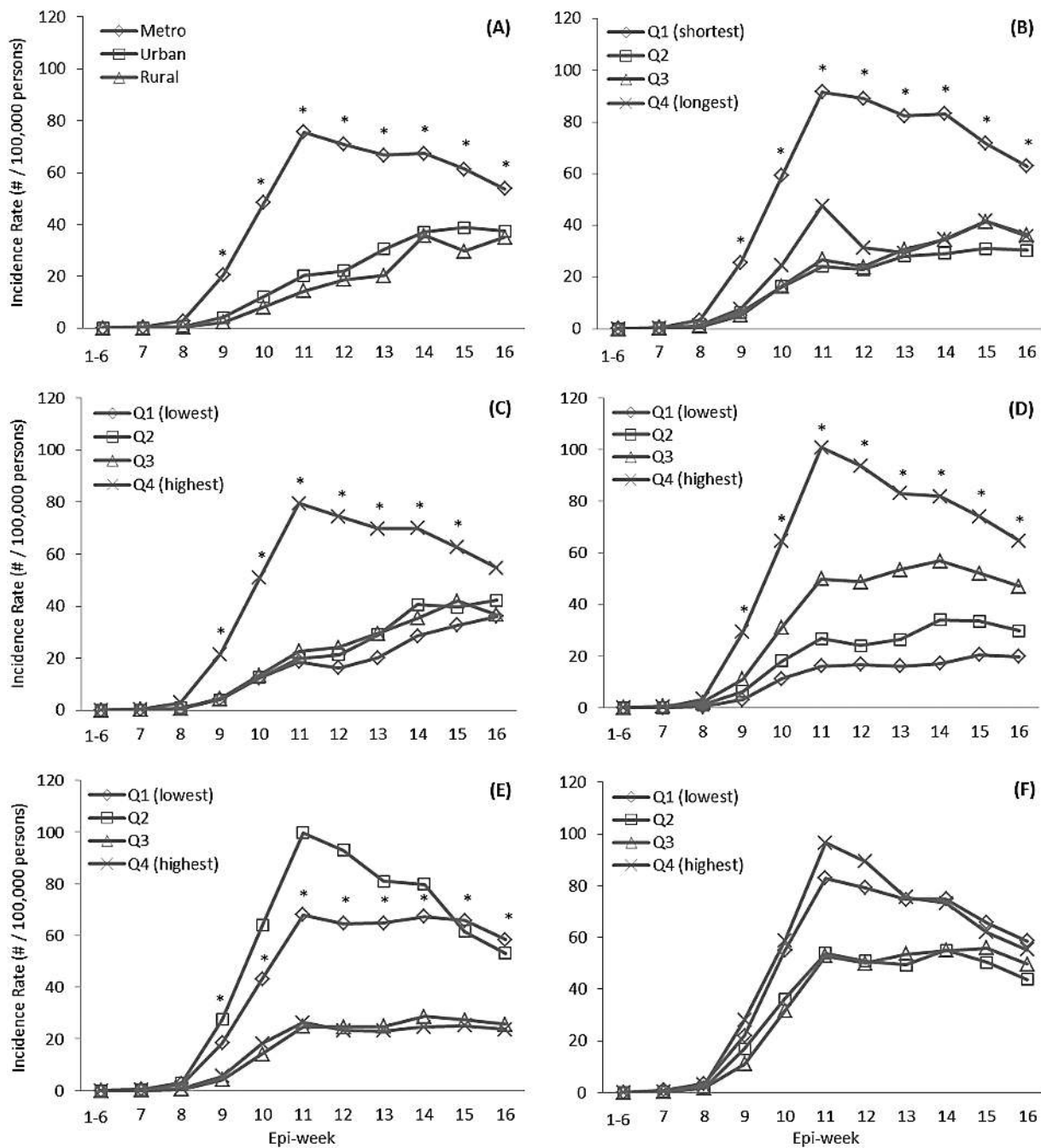
Figure 3. Weekly incidence rates of COVID-19 during 11 observation periods (1–6 weeks in combination and each of the 7–16 weeks), 22 January–13 May 2020, by characteristics of counties in the United States, including rurality (*A*), distances to nearest core airports (*B*), population density (*C*), percentage of minority population (*D*), percentage of people 65 years or older (*E*), and poverty (*F*). *P < .05 for the comparison: metropolitan versus rural, shortest versus longest distance to nearest core airports, and the highest versus lowest shortest versus the longest distance to nearest core airports. Abbreviations: COVID-19, coronavirus disease 2019; Q, quartile.

in southern and northern California, and then rapidly expanded nationwide. Higher risks of COVID-19 clustering and incidence were observed in metropolitan versus rural counties, counties closest to core airports, the most populous counties, and counties with the highest proportion of racial/ethnic minorities. However, the differences have decreased since the 11th epi-week, which was driven by a significant decrease in

the incidence in these counties and a consistent increase in other areas in the recent 5 weeks. This might be a result of social-distancing measures well implemented recently in high-risk areas in the early stage of the outbreak, and also suggests that recent region-to-region spread and community transmission occurred in other areas. Further studies are needed to assess the effectiveness of public health and behavioral

interventions on SARS-CoV-2 infection and implemental barriers, which is essential for promoting the strict adherence to social-distancing guidelines and enhancing personal protections (including appropriately wearing face masks as needed and timely handwashing) to prevent the spread of SARS-CoV-2 and thus substantially decreasing the incidence of COVID-19 locally and nationwide.

A significant association between short distance to core airports and COVID-19 clustering suggests a critical role of air transportation in the spread of SARS-CoV-2 across the country. Air transportation was believed to accelerate and amplify the spread of influenza, SARS-CoV, or MERS-CoV (Middle East respiratory syndrome coronavirus) [13]. A recent study showed that rail transport is related to the transmission and regional outbreak of COVID-19 in China [14]. In the United States, the airports may have contributed substantially to the early travel-related region-to-region transmission. From 18 March to 22 April, at least 42 states, 3 counties, 10 cities, the District of Columbia, and Puerto Rico joined Illinois, New York, and California in the lockdown orders [15]. However, airlines, as one of the essential transportation services, are generally exempted from the orders and are still operating. We flagged the importance of airports in spreading COVID-19 even after the lockdown of most regions in mid-March. Airlines have put in place stringent safeguards for those still flying, including supercharged cleaning, reduced in-flight services, and the spacing out of passengers on flights. It is crucial to maintain strict management and monitoring of major airports to maximize the reduction in region-to-region transmission.

While COVID-19 incidence in metropolitan areas has decreased since the 11th epi-week, we identified a consistent increase in the incidence of COVID-19 in rural areas over the 16 epi-weeks. This was probably a sign that the local spread of COVID-19 extended beyond metro/urban enclaves and secondary community transmission took place around geographic hotspots and spread to rural areas. Rural areas with a lower population density are not safe in this pandemic because rural residents tend to be older and have limited access to healthcare [16, 17]. Therefore, restrictive social-distancing measures are necessary in rural areas, and adherence to social distancing should be enhanced for rural residents.

The pandemic of COVID-19 poses different challenges for US states currently designing their coping strategies. Population density is a key driver for transmission of infectious disease. We observed that predominantly minority counties were at higher risk of COVID-19. This was consistent with the reports of African Americans accounting for about 70% of COVID-19–related deaths but just approximately 30% of the population in Chicago, Milwaukee County, and Louisiana [18]. The disproportionate burden of COVID-19 in minority populations may largely result from inequities in adherence to social-distancing

measures. Our analysis indicates that the incidence of COVID-19 was lower in areas with a higher percentage of elderly people. This could partly result from the lower mobility of older versus younger people. However, we found a comparable risk of COVID-19 clustering in counties with the highest versus lowest percentage of elderly people in recent weeks, which might be relevant to the outbreaks of COVID-19 in nursing homes in some geographic areas [19, 20]. Therefore, areas with a large elderly population should not be ignored in the allocation of prevention efforts on COVID-19 because elderly individuals typically have multiple chronic health conditions and a higher risk of developing more serious complications from COVID-19 [21]. A lack of a stable pattern in the association between poverty and the risk of COVID-19 indicates that socioeconomic factors might not play a critical role in the risk of SARS-CoV-2 infection.

This study has some limitations. The confirmed cases of COVID-19 might not reflect the actual number of persons infected with SARS-CoV-2 due to unknown/untested asymptomatic cases [22–24]. We used reliable governmental records of laboratory-confirmed cases of COVID-19 in the first 16 epi-weeks of the outbreak. Due to limited testing resources, patients with symptoms of COVID-19 were given priority for testing at the early stages of the outbreak. The results of this study might mainly reflect spatiotemporal characteristics of laboratory-confirmed symptomatic cases of COVID-19. Studies to investigate the association of local testing capacity and eligible criteria with spatiotemporal characteristics of COVID-19 are warranted.

This study demonstrated the spatiotemporal characteristics and trends of COVID-19 in the United States, which is essential to better focus our preventive efforts on COVID-19. A critical role of the airports in COVID-19 transmission and clustering suggests that national management and monitoring of major operating airports should be maintained to decrease the risk of region-to-region transmission. At the local levels, social-distancing measures have been widely implemented by cancelling all public events and shuttering all public places [25–28]. The Centers for Disease Control and Prevention (CDC) has launched a public education campaign to promote handwashing and self-hygiene [25–28]. Many cities are disinfecting subways and buses and reducing the metro services [25–28]. Our results suggest that public health guidelines should be further enhanced in high-risk areas of COVID-19 and behavioral interventions are necessary to substantially promote strict adherence to social-distancing guidelines and individual protections. Future research should identify multilevel strategies to enforce behavioral guidelines for COVID-19 prevention, particularly in high-risk geographic areas. The increasing trend of the risk of SARS-CoV-2 infection in rural areas indicates the public health challenges not only in access to healthcare services but also in trust in medical system and access to health insurance

among rural residents. Multilevel public health interventions from government agencies/organizations could be the most effective strategy in controlling and preventing the long-distance spread and local community transmission of SARS-CoV-2 [29–31]. A full understanding of the spatiotemporal dynamics and trends of the COVID-19 epidemic is essential to plan further public health interventions for substantially decreasing the incidence, and to inform the decision making with regard to the timeline of re-opening businesses and public areas through a real-time local risk assessment [32].

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

References

1. World Health Organization. Pneumonia of unknown cause—China. Disease outbreak news. 5 January **2020**. Available at: https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/. Accessed 15 May 2020.
2. Zhu N, Zhang D, Wang W, et al; China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med **2020**; 382:727–33.
3. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis **2020**; 20:533–4.
4. Holshue ML, DeBolt C, Lindquist S, et al; Washington State 2019-nCoV Case Investigation Team. First case of 2019 novel coronavirus in the United States. N Engl J Med **2020**; 382:929–36.
5. USAFacts. Coronavirus stats & data. **2020**. Available at: https://usafacts.org/issues/coronavirus/. Accessed 14 May 2020.
6. Centers for Disease Control and Prevention. Cases of coronavirus disease (COVID-19) in the U.S. **2020**. Available at: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html. Accessed 15 May 2020.
7. The New York Times. Coronavirus in the U.S.: latest map and case count. **2020**. Available at: https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html. Accessed 15 May 2020.
8. Meilleur A, Subramanian SV, Plascak JJ, Fisher JL, Paskett ED, Lamont EB. Rural residence and cancer outcomes in the United States: issues and challenges. Cancer Epidemiol Biomarkers Prev **2013**; 22:1657–67.
9. US Government Accountability Office. Air traffic control modernization: progress and challenges in implementing NextGen. **2017**. Available at: https://www.gao.gov/assets/690/687072.pdf. Accessed 15 May 2020.
10. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. Stat Med **1995**; 14:799–810.
11. Kulldorff M, Heffernan R, Hartman J, Assunção R, Mostashari F. A space-time permutation scan statistic for disease outbreak detection. PLoS Med **2005**; 2:e59.
12. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for Joinpoint regression with applications to cancer rates. Stat Med **2000**; 19:335–51.
13. Browne A, Ahmad SS, Beck CR, Nguyen-Van-Tam JS. The roles of transportation and transportation hubs in the propagation of influenza and coronaviruses: a systematic review. J Travel Med **2016**; 23:tav002.
14. Zhao S, Zhuang Z, Ran J, et al. The association between domestic train transportation and novel coronavirus (2019-nCoV) outbreak in China from 2019 to 2020: a data-driven correlational report. Travel Med Infect Dis **2020**; 33:101568.
15. The New York Times. See which states and cities have told residents to stay at home. **2020**. Available at: https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html. Accessed 15 May 2020.
16. Bull CN, Krout JA, Rathbone-McCuan E, Shreffler MJ. Access and issues of equity in remote/rural areas. J Rural Health **2001**; 17:356–9.
17. Larson SL, Fleishman JA. Rural-urban differences in usual source of care and ambulatory service use: analyses of national data using Urban Influence Codes. Med Care **2003**; 41:III65–74.
18. Yancy CW. COVID-19 and African Americans. JAMA **2020**; 323:1891–2.
19. Grabowski DC, Mor V. Nursing home care in crisis in the wake of COVID-19. JAMA **2020**; 324:23–4.
20. Werner RM, Hoffman AK, Coe NB. Long-term care policy after Covid-19—solving the nursing home crisis. N Engl J Med **2020**. doi: 10.1056/NEJMp2014811
21. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet **2020**; 395:1054–62.
22. Rothe C, Schunk M, Sothmann P, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. N Engl J Med **2020**; 382:970–1.
23. Pan X, Chen D, Xia Y, et al. Asymptomatic cases in a family cluster with SARS-CoV-2 infection. Lancet Infect Dis **2020**; 20:410–11.
24. Day M. Covid-19: four fifths of cases are asymptomatic, China figures indicate. BMJ **2020**; 369:m1375.
25. Centers for Disease Control and Prevention. Social distancing, quarantine, and isolation: keep your distance to slow the spread. **2020**. Available at: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html. Accessed 15 May 2020.
26. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? Lancet **2020**; 395:931–4.
27. Garg S, Kim L, Whitaker M, et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019—COVID-NET, 14 states, March 1–30, 2020. Morb Mortal Wkly Rep **2020**; 69:458–64.
28. Kupferschmidt K, Cohen J. Can China's COVID-19 strategy work elsewhere? Science **2020**; 367:1061–2.
29. Gostin LO, Wiley LF. Governmental public health powers during the COVID-19 pandemic: stay-at-home orders, business closures, and travel restrictions. JAMA **2020**; 323:2137–8.
30. Pan A, Liu L, Wang C, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. JAMA **2020**; 323:1915–23.
31. Hartley DM, Perencevich EN. Public health interventions for COVID-19: emerging evidence and implications for an evolving public health crisis. JAMA **2020**; 323:1908–9.
32. Angulo FJ, Finelli L, Swerdlow DL. Reopening society and the need for real-time assessment of COVID-19 at the community level. JAMA **2020**;3232247–8.