



Association between population density and infection rate suggests the importance of social distancing and travel restriction in reducing the COVID-19 pandemic

Heliang Yin^{1,2,3} · Tong Sun^{3,4} · Lan Yao⁵ · Yan Jiao² · Li Ma^{1,2,3} · Lin Lin^{1,2,3} · J Carolyn Graff⁶ · Lotfi Aleya⁷ · Arnold Postlethwaite^{8,9} · Weikuan Gu^{2,9}  · Hong Chen^{1,3}

Received: 8 August 2020 / Accepted: 2 January 2021 / Published online: 13 January 2021

© This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2021

Abstract

Currently, 2019-nCoV has spread to most countries of the world. Understanding the environmental factors that affect the spread of the disease COVID-19 infection is critical to stop the spread of the disease. The purpose of this study is to investigate whether population density is associated with the infection rate of the COVID-19. We collected data from official webpages of cities in China and in the USA. The data were organized on Excel spreadsheets for statistical analyses. We calculated the morbidity and population density of cities and regions in these two countries. We then examined the relationship between morbidity and other factors. Our analysis indicated that the population density in cities in Hubei province where the COVID-19 was severe was associated with a higher percentage of morbidity, with an r value of 0.62. Similarly, in the USA, the density of 51 states and territories is also associated with morbidity from COVID-19 with an r value of 0.55. In contrast, as a control group, there is no association between the morbidity and population density in 33 other regions of China, where the COVID-19 epidemic is well under control. Interestingly, our study also indicated that these associations were not influenced by the first case of COVID-19. The rate of morbidity and the number of days from the first case in the USA have no association, with an r value of -0.1288 . Population density is positively associated with the percentage of patients with COVID-19 infection in the population. Our data support the importance of such as social distancing and travel restriction in the prevention of COVID-19 spread.

Keywords COVID-19 · Density · Population · Social distancing · Transmission

Heliang Yin and Tong Sun contributed equally to this work.

Responsible Editor: Philippe Garrigues

✉ Weikuan Gu
wgu@uthsc.edu

✉ Hong Chen
qszywdm@163.com

¹ Center of Integrative Research, The First Hospital of Qiqihar City, 30 Gongyuan Road, Qiqihar, Heilongjiang 161005, People's Republic of China

² Department of Orthopedic Surgery and BME-Campbell Clinic, University of Tennessee Health Science Center, 956 Court Avenue, Memphis, TN 38163, USA

³ Affiliated Qiqihar Hospital, Southern Medical University, Qiqihar, Heilongjiang 161007, People's Republic of China

⁴ Department of Administration, The First Hospital of Qiqihar, Qiqihar, Heilongjiang 161005, People's Republic of China

⁵ Health Outcomes and Policy Research, College of Graduate Health Sciences, University of Tennessee Health Science Center, Memphis, TN 38103, USA

⁶ College of Nursing, University of Tennessee Health Science Center, Memphis, TN 38105, USA

⁷ Chrono-Environnement Laboratory, UMR CNRS 6249, Bourgogne Franche-Comté University, F-25030 Besançon Cedex, France

⁸ Department of Medicine and Division of Connective Tissue Diseases, University of Tennessee Health Science Center, Memphis, TN 38163, USA

⁹ Research Service, Memphis VA Medical Center, 1030 Jefferson Avenue, Memphis, TN 38104, USA

Introduction

At present, there is a worldwide pandemic of COVID-19 (Wang et al. 2020a). With its associated morbidity and mortality, COVID-19 is on the track to become one of the most catastrophic pandemics in human history (Wang et al. 2020a). A considerable amount of research and publications have focused on the analysis of the factors that lead to COVID-19 infection (Koo et al. 2020; Sun et al. 2020). After its person-to-person transmission was confirmed, the influence of environmental factors on COVID-19 transmission in human populations has received considerable attention. Several environmental factors, such as the nature of surfaces of objects and the role of transmission airborne infections, have attracted great attention from the public. Further systematic understanding of the impact of environmental factors on human-to-human transmission of this virus may be extremely important in designing measures to contain the COVID-19 epidemic (Bedford et al. 2020; Jarvis et al. 2020; Lasry et al. 2020., Nakada and Urban 2020). Among many environmental factors, population density is one of the conditions that cannot be underestimated and may affect the infection rate of COVID-19. Population density may directly reflect on whether and how social distancing and travel restriction work to slow the spread of COVID-19 (Gewin 2020; Gibson and Rush 2020). In theory, the higher the population density, the greater the chance of COVID-19 infection. However, it is still unclear to what extent population density affects the infection rate of COVID-19.

The current available data in coronavirus infections provides an opportunity to conduct a preliminary analysis of the impact of population density on the infection rate of COVID-19. For example, the number of people infected in different cities with different population densities in China has been reported (Chen et al. 2020; Huang et al. 2020). COVID-19 infection rates and mortality rates in different states and territories of the USA are currently reported and updated daily (Lasry et al. 2020). Here we present a preliminary analysis of the impact of population density on the rate of COVID-19 infections.

Method

Data collection from different regions and cities located in China and the USA

Data collection was from three sets of multiple locations. The first set of locations is the 17 cities in the Hubei province in China. These cities are all located close to Wuhan and had large numbers of COVID-19 patients among cities in China. Data from these cities were from the official website of Hubei province. Next, we collected second set of data from other

cities and provinces of China, hereafter referred to as other regions. The differences between these other regions and cities in Hubei are that less people contracted COVID-19 in other regions than in the cities of the Hubei province. COVID-19 disease in these other regions was well tracked, and disease incidence did not reach the pandemic level. The data from these other regions are all from provincial websites (<http://wjw.hubei.gov.cn/fbjd/dtyw/>). The third set of data is from the 50 states and three territories of the USA. These data were obtained from a website (<https://www.cdc.gov/coronavirus/2019-nCoV/index.html>). Data collected include the names of the city/state/ territory, accumulative deaths, population size in (millions), and geographical area in (square kilometers).

Data organization and calculation

Data were uploaded and organized using the Excel software. The population density of a city was calculated by dividing the population of the city by the geographical area. Thus, the result was the number of persons per square kilometer. The infection morbidity rate of a given location was obtained by dividing the number of COVID-19 cases by the population of the city. The relationship between the population density and the disease morbidity then was analyzed by correlation coefficients. Linear regression model was used to demonstrate their mathematic relationship.

First cases in the USA and territories

In order to examine the relationship between the disease morbidity and the time of the first case in a city, we collected the information on the first cases in these cities and coded their first cases for the analysis. Based on dates of reported onsets of the first cases in the US states and territories, we assigned numbers based on the date of the first case. We first assigned the earliest US case in Washington State in Seattle on January 21, 2020 as #1. The number assigned to the other locations is derived by adding 1 to the difference between date of the first case in a given city and the day of first case in Seattle, January 21, 2020. For example, the first case reported in Arizona was on January 26, 2020. The Arizona number therefore is $1 + (\text{January } 26 - \text{January } 21) = 6$.

Statistical analyses

For the relationship analysis, we used the previous criteria to categorize the strength of the Pearson correlation (Wang et al. 2020b). Thus, if the r value is greater than or equal to 0.7 or is less than or equal to -0.7 , it is a strong positive or negative correlation, respectively. If the r value is between 0.35 and 0.69 or -0.35 and -0.69 , a correlation exists. However, r

values falling between 0 and 0.35 or 0 and -0.35 were regarded as not correlated.

Results

Basic patterns and information from different regions

Our sets of data include cities of Hubei province, other regions in China, and 53 states and territories in the USA. The first set of data was obtained from 17 cities in the Hubei Province (Supplemental Table 1). These 17 cities had a total of 68,128 patients diagnosed with COVID-19 and 4512 deaths from COVID-19. The total population in these 17 cities is 59.965 million, living in an area of 162,245 km². The calculated average COVID-19 mortality rate is 0.665‰. The second set of data is the population density and the total number of people in other provinces and cities or other regions in China (Supplemental Table 2). The collected data show that the average population density of these cities and provinces is 1106 km². However, the morbidity rate in these places is very low, and the prevalence rate is 0.025 per thousand. The third set of data is the population density and prevalence of 53 major cities in states and territories in the USA (Supplemental Table 3). The population density of these cities is averaged to be 90 per square kilometer. The morbidity is 1.412 per thousand people.

Association between population density and mortality in Hubei province

Our data indicate that population density of the 17 cities in the Hubei province is positively associated with the disease mortality. There are considerable differences in the population density and death rate among these 17 cities. Although on average, the population density is 489 persons per square kilometers, the difference from city to city is large, ranging from 24.6 in Shennongjia to 1461 in Enshi (Fig. 1a). Similarly, on average, the average disease morbidity is rate is 3.49% ranging from 0.025 to 4.167% (Fig. 1b). Correlation analysis indicated that there is a positive association between the population density and disease morbidity, with an r value of 0.620 (Fig. 1c).

Association between population density and morbidity in major cities in the US states and territories. Our data indicate that population density in the US states and territories is positively associated with disease morbidity. The population density of the US states and territories is much lower than that of China; nevertheless, there are also considerable differences in the density and death rate among different locations in the US states and territories. On average, the US population density is 80 persons per square kilometer; the

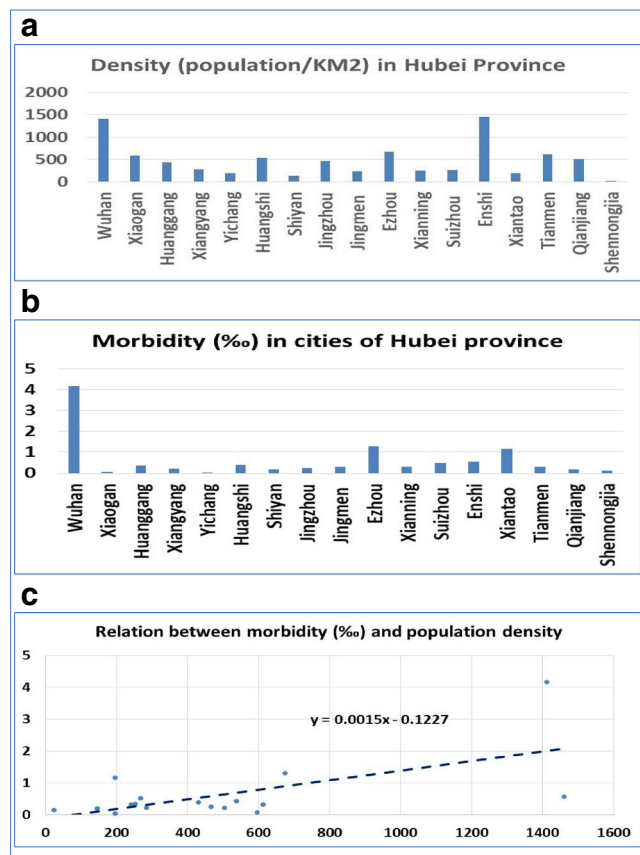
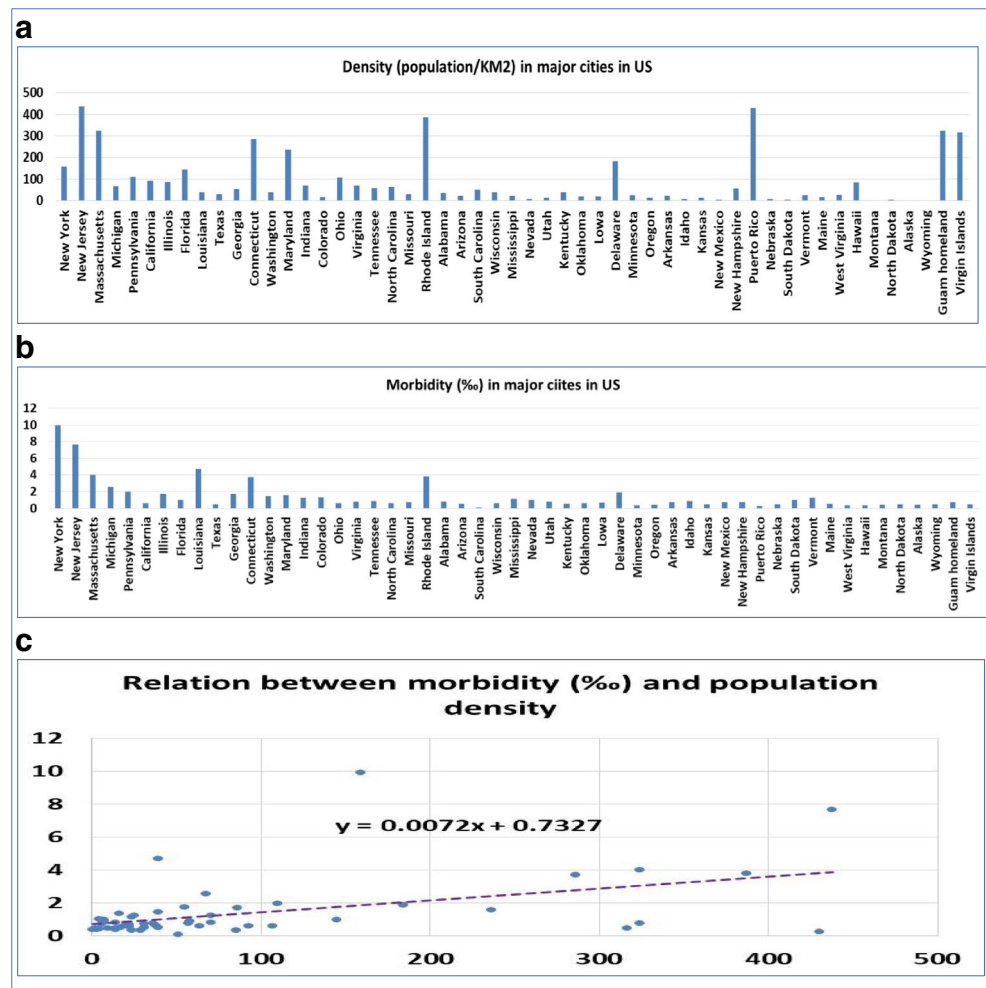


Fig. 1 Population density and morbidity in Hubei province. **a** Population density in the 17 cities in Hubei province. The number on the vertical axis is the average number of persons per kilometer. Names of the cities are listed on the horizontal axis. **b** Disease morbidity rate in the 17 cities in Hubei province. The number on vertical axis is the average rate of morbidity. Names of the cities are listed on the horizontal axis. **c** The relationship between the population density and disease morbidity analyzed by simple linear regression. The number on vertical axis is the disease morbidity rate, and the horizontal axis is the number of persons per kilometer

difference from state to state is from as low as 0.49 in Alaska to as high as 438.00 in New Jersey (Fig. 2a). Similarly, on average, the disease morbidity has a rate of 1.41%, with a variation from 0.07% in South Carolina to 9.94% in New York (Fig. 2b). Further comparison indicates that there is a positive correlation between the population density and disease morbidity, with an r value of 0.552 (Fig. 2c). There is apparently a highly consistent relationship between population density and COVID-19 mortality.

Because the date of the first cases reported in different cities in the USA is considerably different, and the disease epidemic is still ongoing, we analyzed the relation between the date of the first case and the morbidity to determine whether the date of the first case influenced the rate of morbidity (Supplemental Table 4). Surprisingly, there was no association between them, with an r value of -0.1288 .

Fig. 2 Population density and morbidity in 53 locations in the USA and its territories. **a** Population density in the 53 locations in the US states and territories. The number on vertical axis is the average number of persons per kilometer. Names of the states and territories are listed on the horizontal axis. **b** Disease morbidity rate in the 53 in the US states and territories. The number on vertical axis is the average rate of morbidity from COVID-19. Names of the states and territories are listed on the horizontal axis. **c** The relationship between the population density and disease morbidity. The number on vertical axis is the disease morbidity rate, and the horizontal axis is the number of persons per kilometer



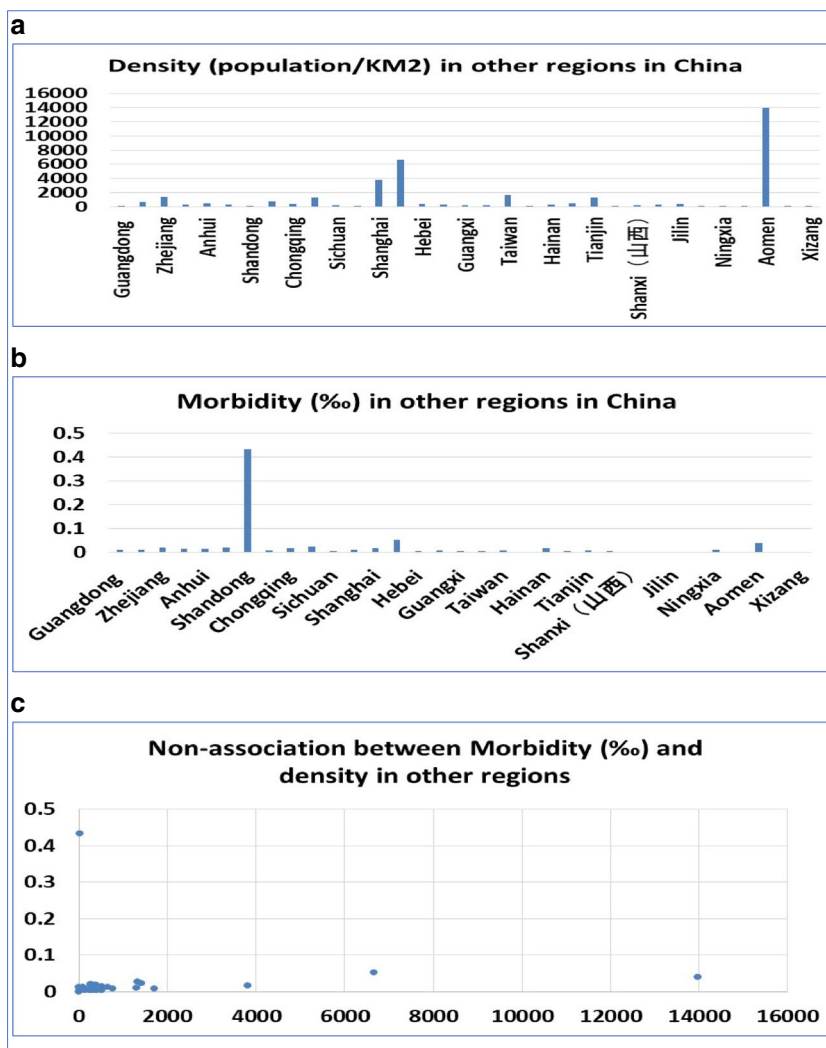
Non-association between population density and mortality among other regions in China

Since the disease in other regions of China was well under control, an epidemic did not materialize into the general population before the elimination of the infection source. We hypothesized that in this case, the population density in other regions would not be associated with the disease morbidity. These regions include 33 regions, with an average of 1106 persons per kilometer but with large differences from 2.80 persons per kilometer in XiZang to 13,984 persons per kilometer in Aomen (Fig. 3a). The disease morbidity rate is low, all below 0.1%, except Shandong which had a rate of 0.4% (Fig. 3b). The *r* value for the correlation between population density and disease morbidity rate is 0.04 (Fig. 3c). Thus, the data from other regions in China serve as a negative control; when the COVID-19 disease does not morph into an epidemic, population density is not associated with the disease morbidity.

Difference between China and the USA on the measures of social distance and the impact of disease epidemic

By analyses of the data from the Hubei province in China and 53 major US cities in the US states and territories, we obtained the positive correlations between the population density and the disease morbidity. We next asked whether there is a difference between these two data sets of cities on the impact of the disease epidemic. We calculated the morbidity of population density from 100 from 1000 using the linear formula obtained from the Hubei and major cities of the US states and territories. Although both data sets have positive correlation between the population density and morbidity, they are not at the same degree (Table 1). Based on the formula derived from Hubei province, the morbidity increased from 0.177 to 1.377 when the population density increased from the 100 to the 1000 (Fig. 4a). On the other hand, in major cities and territories of the USA, the morbidities increased from 1.453 to 7.93 when the population density increased from the 100 to

Fig. 3 Non-association between population density and morbidity in other regions in China. **a** Population density in the rest 33 regions in China. The number on vertical axis is the average number of persons per kilometer. Names of the cities are listed on the horizontal axis. **b** Disease morbidity rate in the 33 regions in China. The number on the vertical axis is the average rate of morbidity. Names of the cities are listed on the horizontal axis. **c** The relationship between the population density and disease morbidity. The number on vertical axis is the disease morbidity rate, and the horizontal axis is the number of persons per kilometer



the 1000 (Fig. 4b). The increases in rates between these two sets of data are significantly different (Fig. 4c).

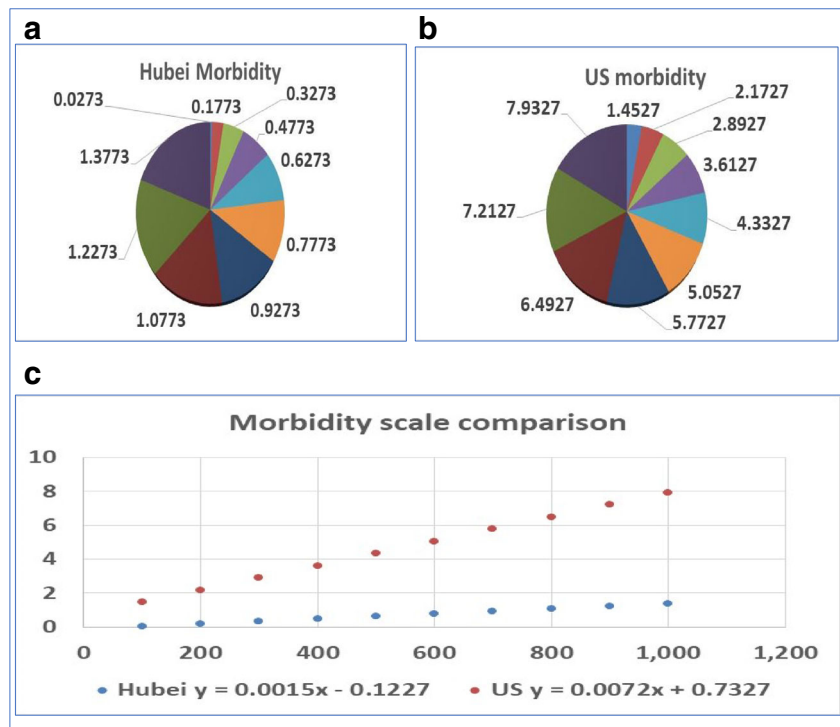
Table 1 Predicated morbidity at different population density in the USA and Hubei, China

Population density	US morbidity	Population density	Hubei morbidity
100	1.4527	100	0.0273
200	2.1727	200	0.1773
300	2.8927	300	0.3273
400	3.6127	400	0.4773
500	4.3327	500	0.6273
600	5.0527	600	0.7773
700	5.7727	700	0.9273
800	6.4927	800	1.0773
900	7.2127	900	1.2273
1000	7.9327	1000	1.3773

Discussion

Our analyses clarified the positive correlation between population density and the epidemic infection rate. From the data of 17 cities in the Hubei province, the correlation between population density and the number of people infected with COVID-19 in this epidemic is 0.62, showing a relatively high positive correlation. An analysis of 53 locations in the US states and territories yielded a correlation of 0.55 when examining the relation between population density and the number of people previously infected. Additionally, the number of infected patients is not directly related to the date of initial infection. These two sets of data show that due to the higher population density, the chance of contact between people is relatively high. When the coefficient representing the relation between population density and number of persons infected is relatively high, the absolute number of people who contact each other is relatively high. The findings from this study indirectly support the practice of keeping social distance and enforcing of travel restriction (Vuong et al. 2020).

Fig. 4 Comparison of the morbidity scales between major cities in the Hubei province of China and the US states and territories. **a** The morbidities at ten different scales were calculated based on the formula derived from 17 cities of the Hubei province in China. **b** The morbidities at ten different scales were calculated based on the formula derived from 53 major cities of the US states and territories. **c** The comparison of disease morbidities between cities in China and in the US states and territories when the population density increases from 100 to 1000 persons per kilometer. The number on the vertical axis is the disease morbidity rate, and the horizontal axis is the number of persons per kilometer



Our data showed that given the same population density, the scale of disease morbidity in the US states and territories is much higher than that in China. Given the complex differences between the social and economic systems and ethnic populations in these two countries, the reason for such a difference needs to be explored in a much broader investigation. Our study only reveals that such a difference exists but does not explain why there is such a difference.

In addition, this study also shows that in the case when the epidemic is well controlled, the initial source of infection is detected early and the route of infection could be cut off. In the control set, when there is no major outbreak, the population density is not related to the number of people infected. Thus, in other regions in China, regardless of density in populations, COVID-19 has not had a major outbreak. This is supported by the finding that the correlation coefficient measuring the relation between population density and population infection in the other regions in China is 0.037.

We realized that due to the back and forth revision and correction of the data by the official sources, it is not possible that all the data are error-free. However, these data as a whole are reliable.

In summary, our data show that the extent and speed of a large outbreak under the same conditions are directly affected by the population density, which is of great significance to the government’s decision-making on social distancing and travel restrictions when responding to the outbreak.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11356-021-12364-4>.

Authors’ contributions HC, WG, YJ, LA, HC, and DS conceived and designed the experiments. TS, HY, LY, and WG performed data searching and collection. TS, HY, LY, LM, LL, and WG analyzed the data. WG, CH, HC, and DS contributed analysis tools. TS, HY, LY, LA, SD, CG, SH, and WG wrote the manuscript. All authors revise and approve and manuscript.

Funding This work was partially supported by funding from merit grant I01 BX000671 to WG from the Department of Veterans Affairs and the Veterans Administration Medical Center in Memphis, TN, USA, and grant 90DDUC0058 to CG from the US Department of Health and Human Services, Administration for Community Living.

Data availability All the data in this study are available for public. Sources of the data have been provided in the text and references of this publication.

Compliance with ethical standards

Competing interests The authors declare that they have no competing interests.

Ethical approval N/A

Consent to participate N/A

Consent to publish N/A

References

Bedford J, Enria D, Giesecke J, Heymann DL, Ihekweazu C, Kobinger G, Lane HC, Memish Z, Oh MD, Sall AA, Schuchat A, Ungchusak K, Wieler LH (2020) WHO strategic and technical advisory group for

- infectious hazards. COVID-19: towards controlling of a pandemic. *Lancet*. 395(10229):1015–1018
- Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 395:507–513
- Gewin V (2020 Apr) Safely conducting essential research in the face of COVID-19. *Nature*. 580(7804):549–550
- Gibson L, Rush D (2020 Apr 6) Novel coronavirus in Cape Town informal settlements: feasibility of using informal dwelling outlines to identify high risk areas for COVID-19 transmission from a social distancing perspective. *JMIR Public Health Surveill*. 6(2):e18844
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395(10223):497–506
- Jarvis CI, Van Zandvoort K, Gimma A, Prem K, CMMID COVID-19 working group, Klepac P, Rubin GJ, Edmunds WJ (2020) Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med*. 18(1):124
- Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, Tam C, Dickens BL (2020) Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *Lancet Infect Dis* 20:678–688
- Lasry A, Kidder D, Hast M, Poovey J, Sunshine G, Winglee K, Zviedrite N, Ahmed F, Ethier KA; CDC Public Health Law Program; New York City Department of Health and Mental Hygiene; Louisiana Department of Health; Public Health – Seattle & King County; San Francisco COVID-19 Response Team; Alameda County Public Health Department; San Mateo County Health Department; Marin County Division of Public Health (2020) Timing of community mitigation and changes in reported COVID-19 and community mobility - four U.S. metropolitan areas, February 26–April 1, 2020. *MMWR Morb Mortal Wkly Rep* 69(15):451–457
- Nakada LYK, Urban RC (2020 Sep 28) COVID-19 pandemic: environmental and social factors influencing the spread of SARS-CoV-2 in Sao Paulo, Brazil. *Environ Sci Pollut Res Int*:1–7. <https://doi.org/10.1007/s11356-020-10930-w>
- Sun K, Chen J, Viboud C (2020 Apr) Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *Lancet Digit Health*. 2(4):e201–e208
- Vuong QT, Thang PQ, Park MK, Choi SD (2020 Dec) Effects of the COVID-19 lockdown on criteria air pollutants in the city of Daegu, the epicenter of South Korea's outbreak. *Environ Sci Pollut Res Int*. 27(36):45983–45991. <https://doi.org/10.1007/s11356-020-11360-4>
- Wang C, Horby PW, Hayden FG, Gao GF (2020a) A novel coronavirus outbreak of global health concern. *Lancet* 395(10223):470–473
- Wang L, Li J, Guo S, Xie N, Yao L, Cao Y, Day SW, Howard SC, Graff JC, Gu T, Ji J, Gu W, Sun D (2020b Apr 8) Real-time estimation and prediction of mortality caused by COVID-19 with patient information based algorithm. *Sci Total Environ*. 727:138394

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.