

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

3
4 Heliang Yin, MD, MS^{1,2,3*}; Tong Sun, MS^{3, 4*}; Lan Yao, MS^{5, 6}; Yan Jiao, MS²; Li Ma, MD,
5 MS^{1,2,3}; Lin Lin, MD, MS^{1,2,3}; J Carolyn Graff, PhD, RN, FAAIDD⁷; Lotfi Aleya, PhD⁸; Arnold
6 Postlethwaite, MD^{9, 10}; Weikuan Gu, PhD^{2, 10}; Hong Chen MD, PhD^{1, 3}

7
8
9 ¹Center of Integrative Research, The First Hospital of Qiqihar City, Qiqihar, Heilongjiang
10 161005, P.R. China.

11 ²Department of Orthopedic Surgery and BME-Campbell Clinic, University of Tennessee Health
12 Science Center, Memphis, Tennessee, 38163, USA.

13 ³ Affiliated Qiqihar Hospital, Southern Medical University, Qiqihar, Heilongjiang, 161007, P.R.
14 China.

15 ⁴Department of Administration, The First Hospital of Qiqihar, Qiqihar, Heilongjiang, 161005,
16 P.R. China.

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

19 ⁶Center for Endemic Disease Control, Chinese Center for Disease Control and Prevention,
20 Harbin Medical University; Key Laboratory of Etiologic Epidemiology, Education Bureau of
21 Heilongjiang Province & Ministry of Health (23618104), 157 Baojian Road, Harbin,
22 Heilongjiang, 150081, P. R. China.

23 ⁷College of Nursing, University of Tennessee Health Science Center, Memphis, TN 38105,
24 USA.

25 ⁸Chrono-Environnement Laboratory, UMR CNRS 6249, Bourgogne Franche-Comté University,
26 F-25030 Besançon Cedex, France

27 ⁹Department of Medicine and Division of Connective Tissue Diseases, University of Tennessee
28 Health Science Center, Memphis, Tennessee, 38163, USA.

29 ¹⁰Research Service, Memphis VA Medical Center, 1030 Jefferson Avenue, Memphis, TN,
30 38104, USA.

31 *These authors contributed to this manuscript equally.

32

33 Correspondence and requests for materials should be addressed to: Weikuan Gu, 956
34 Court Avenue, Memphis, TN 38163, USA. Tel: 1-901-448-2259; E-mail address:
35 wgu@uthsc.edu; Hong Chen, Center of Integrative Research, The First Hospital of Qiqihar City,
36 30 Gongyuan Road, Qiqihar, Heilongjiang 161005, PR China. Email: qszyywdm@163.com

37

38

39 **Running title:** Population density and infection rate of COVID-19

40

41

42

43

44

45

46

47

48

49

50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77

Abstract

Background: Currently, the 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.

Methods: We collected data from official webpages of cities in China and in the US. The data were uploaded 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 the morbidity and other factors.

Results: 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 US, 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 US has no association, with an r value of -0.1288.

Conclusions: 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.

Key Words: COVID-19; Density; Population; Social distancing; Transmission.

78 **Introduction**

79

80 At present, there is a worldwide pandemic of COVID-19 [Wang, Horby, Hayden, & Gao, 2020].

81 With its associated morbidity and mortality, COVID-19 is on the track to become one of the

82 most catastrophic pandemics in human history [Wang et al., 2020]. A considerable amount of

83 research and publications have focused on the analysis of the factors that lead to COVID-19

84 infection [Koo et al., 2020; Sun et al., 2020]. After its person-to-person transmission was

85 confirmed, the influence of environmental factors on COVID-19 transmission in human

86 populations has received considerable attention. Several environmental factors, such as the

87 nature of surfaces of objects and the role of transmission airborne infections, have attracted great

88 attention from the public. Further systematic understanding of the impact of environmental

89 factors on human-to-human transmission of this virus may be extremely important in designing

90 measures to contain the COVID-19 epidemic [Bedford et al., 2020; Jarvis et al., 2020; Lasry et

91 al., 2020]. Among many environmental factors, population density is one of the conditions that

92 cannot be underestimated and may affect the infection rate of COVID-19. Population density

93 may directly reflect on whether and how social distancing and travel restriction work to slow the

94 spread of COVID-19 [Gewin et al., 2020; Gibson et al., 2020]. In theory, the higher the

95 population density, the greater the chance of COVID-19 infection. However, it is still unclear to

96 what extent population density affects the infection rate of COVID-19.

97 The current available data in coronavirus infections provides an opportunity to conduct a

98 preliminary analysis of the impact of population density on the infection rate of COVID-19. For

99 example, the number of people infected in different cities with different population densities in

100 China has been reported [Chen et al., 2020; Huang et al., 2020]. COVID-19 infection rates and

101 mortality rates in different states and territories of the United States are currently reported and
102 updated daily [Lasry et al., 2020]. Here we present a preliminary analysis of the impact of
103 population density on the rate of COVID-19 infections.

104

105 **Method**

106 1. Data collection from different regions and cities located in China and the US.

107 Data collection were from three sets of multiple locations. The first set of locations is the 17
108 cities in the Hubei province in China. These cities are all located close to Wuhan and had
109 large numbers of COVID-19 patients among cities in China. Data from these cities were
110 from the official website of Hubei province. Next, we collected second set of data from
111 other cities and provinces of China, hereafter referred to as other regions. The differences
112 between these other regions and cities in Hubei is that less people contracted COVID-19 in
113 other regions than in the cities of the Hubei province. COVID-19 disease in these other
114 regions was well tracked, and disease incidence did not reach the pandemic level. The data
115 from these other regions are all from provincial websites
116 (<http://wjw.hubei.gov.cn/fbjd/dtyw/>). The third set of data is from the 50 states and three
117 territories of the US. These data were obtained from a website
118 (<https://www.cdc.gov/coronavirus/2019-nCoV/index.html>). Data collected include the
119 names of the city/state/ territory, accumulative deaths, population size in (millions) and
120 geographical area in (square kilometers).

121 2. Data organization and calculation.

122 Data were uploaded and organized using Excel software. The population density of a city
123 was calculated by dividing the population of the city by the geographical area. Thus, the

124 result was the number of persons per square kilometer. The infection morbidity rate of a
125 given location was obtained by dividing the number of COVID-19 cases by the
126 population of the city. The relationship between the population density and the disease
127 morbidity then was analyzed by correlation coefficients. Linear regression model was
128 used to demonstrate their mathematic relationship.

129 3. First cases in the US and territories

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

139 2. Statistical analyses

140 For the correlation, we used the previous criteria to categorize the strength of the correlation
141 (Wang et al., 2016). Thus, if the r value is greater than or equal to 0.7 or is less than or equal to -
142 0.7, it is a strong positive or negative correlation, respectively. If the r value is between 0.35 and
143 0.69 or -0.35 and -0.69, a correlation exists. However, r values falling between 0 and 0.35 or 0
144 and -0.35, were regarded as not correlated.

145

146 **Results**

147 1. Basic patterns and information from different regions

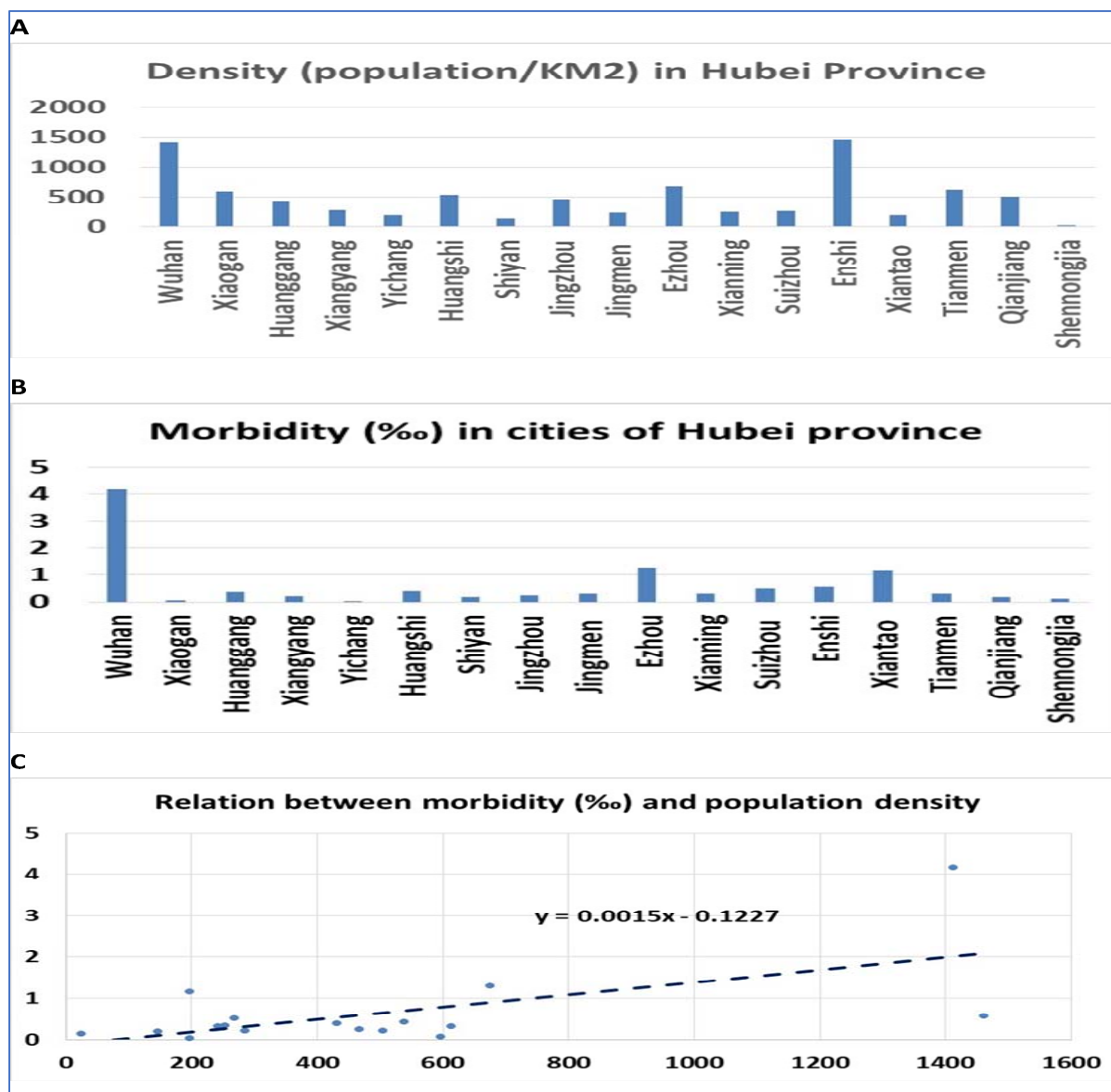
148 Our sets of data include cities of Hubei province, other regions in China, and 53 states and
149 territories in US. The first set of data was obtained from 17 cities in the Hubei Province
150 (Supplemental Table 1). These 17 cities had a total of 68,128 patients diagnosed with
151 COVID-19 and 4,512 deaths from COVID-19. The total population in these 17 cities is
152 59.965 million, living in an area of 162,245 square kilometers. The calculated average
153 COVID-19 mortality rate is 0.665‰. The second set of data is the population density and
154 the total number of people in other provinces and cities or other regions in China
155 (Supplemental Table 2). The collected data show that the average population density of
156 these cities and provinces is 1,106 square kilometers. However, the morbidity rate in these
157 places is very low, and the prevalence rate is 0.025 per thousand. The third set of data is the
158 population density and prevalence of 53 major cities in states and territories in the United
159 States (Supplemental Table 3). The population density of these cities is averaged to be 90
160 per square kilometer. The morbidity is 1.412 per thousand people.

161 2. Association between population density and mortality in Hubei province.

162 Our data indicate that population density of the 17 cities in the Hubei province is positively
163 associated with the disease mortality. There are considerable differences in the population
164 density and death rate among these 17 cities. Although on average, the population density is
165 489 persons per square kilometers, the difference from city to city is large, ranging from
166 24.6 in Shennongjia to 1461 in Enshi (Figure 1A). Similarly, on average, the average
167 disease morbidity is rate is 3.49% ranging from 0.025 to 4.167% (Figure 1B). Correlation
168 analysis indicated that there is a positive association between the population density and
169 disease morbidity, with an r value of 0.620 (Figure 1C).

170

171



172

173 Figure 1. Population density and morbidity in Hubei province. Fig. 1A: Population density

174 in the 17 cities in Hubei province. The number on the vertical axis is the average number of

175 persons per kilometer. Names of the cities are listed on the horizontal axis. Fig. 1B: Disease

176 morbidity rate in the 17 cities in Hubei province. The number on vertical axis is the average

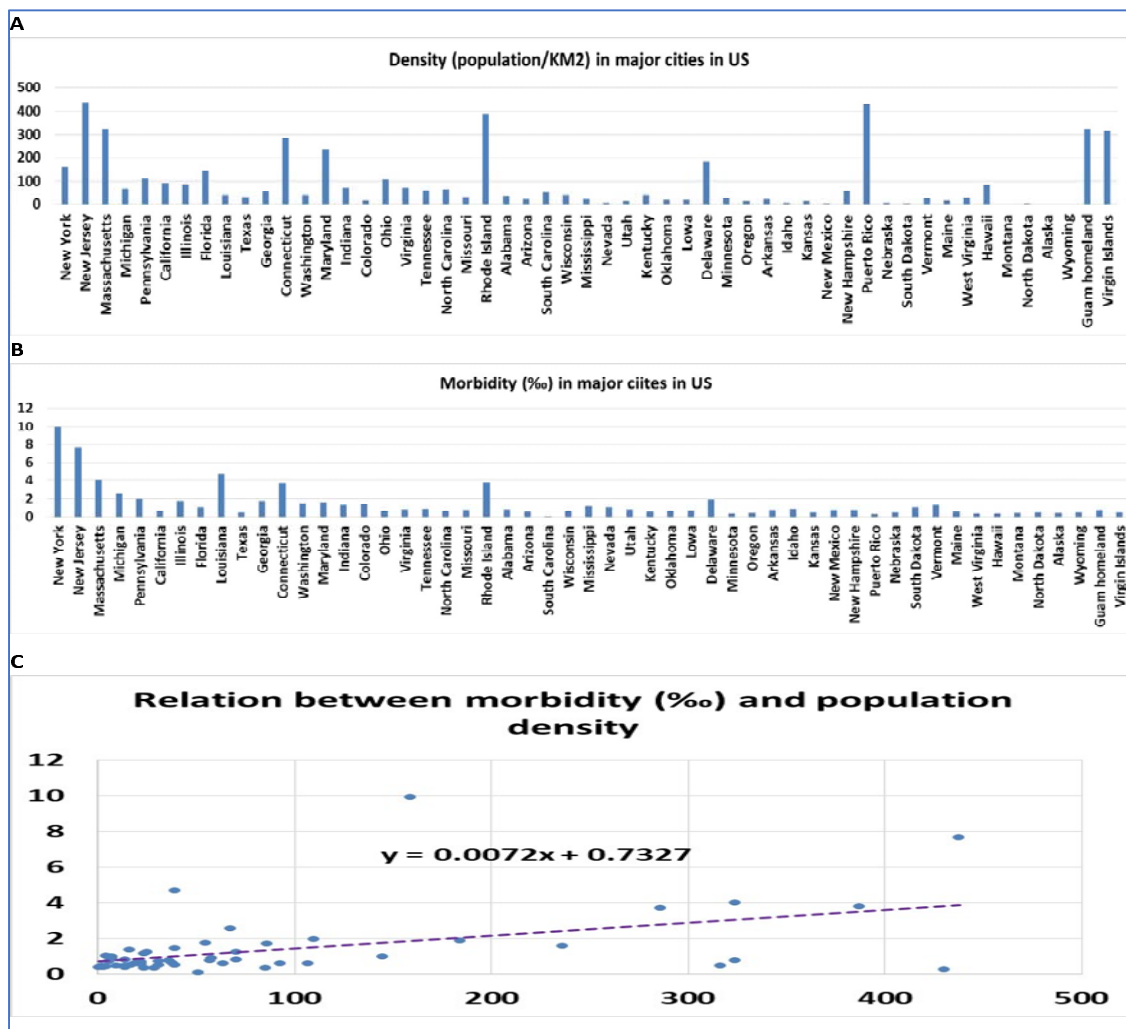
177 rate of morbidity. Names of the cities are listed on the horizontal axis. Fig. 1C. The
178 relationship between the population density and disease morbidity. The number on vertical
179 axis is the disease morbidity rate, and the horizontal axis is the number of persons per
180 kilometer.

1813. Association between population density and morbidity in major cities in the US states and
182 territories. Our data indicate that population density in the US states and territories is positively
183 associated with disease morbidity. The population density of the US states and territories is
184 much lower than that of China; nevertheless, there are also considerable differences in the
185 density and death rate among different locations in the US states and territories. On average, the
186 US population density is 80 persons per square kilometer; the difference from state to state is
187 from as low as 0.49 in Alaska to as high as 438.00 in New Jersey (Figure 2A). Similarly, on
188 average, the disease morbidity has a rate of 1.41%, with a variation from 0.07% in South
189 Carolina to 9.94% in New York (Figure 2B). Further comparison indicates that there is a positive
190 correlation between the population density and disease morbidity, with an r value of 0.552
191 (Figure 2C). There is apparently a highly consistent relationship between population density and
192 COVID-19 mortality.

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

198

199



200

201

202 Figure 2. Population density and morbidity in 53 locations in the US and its territories. Fig.

203 2A: Population density in the 53 locations in US states and territories. The number on

204 vertical axis is the average number of persons per kilometer. Names of the states and

205 territories are listed on the horizontal axis. Fig. 2B: Disease morbidity rate in the 53 in US

206 states and territories. The number on vertical axis is the average rate of morbidity from

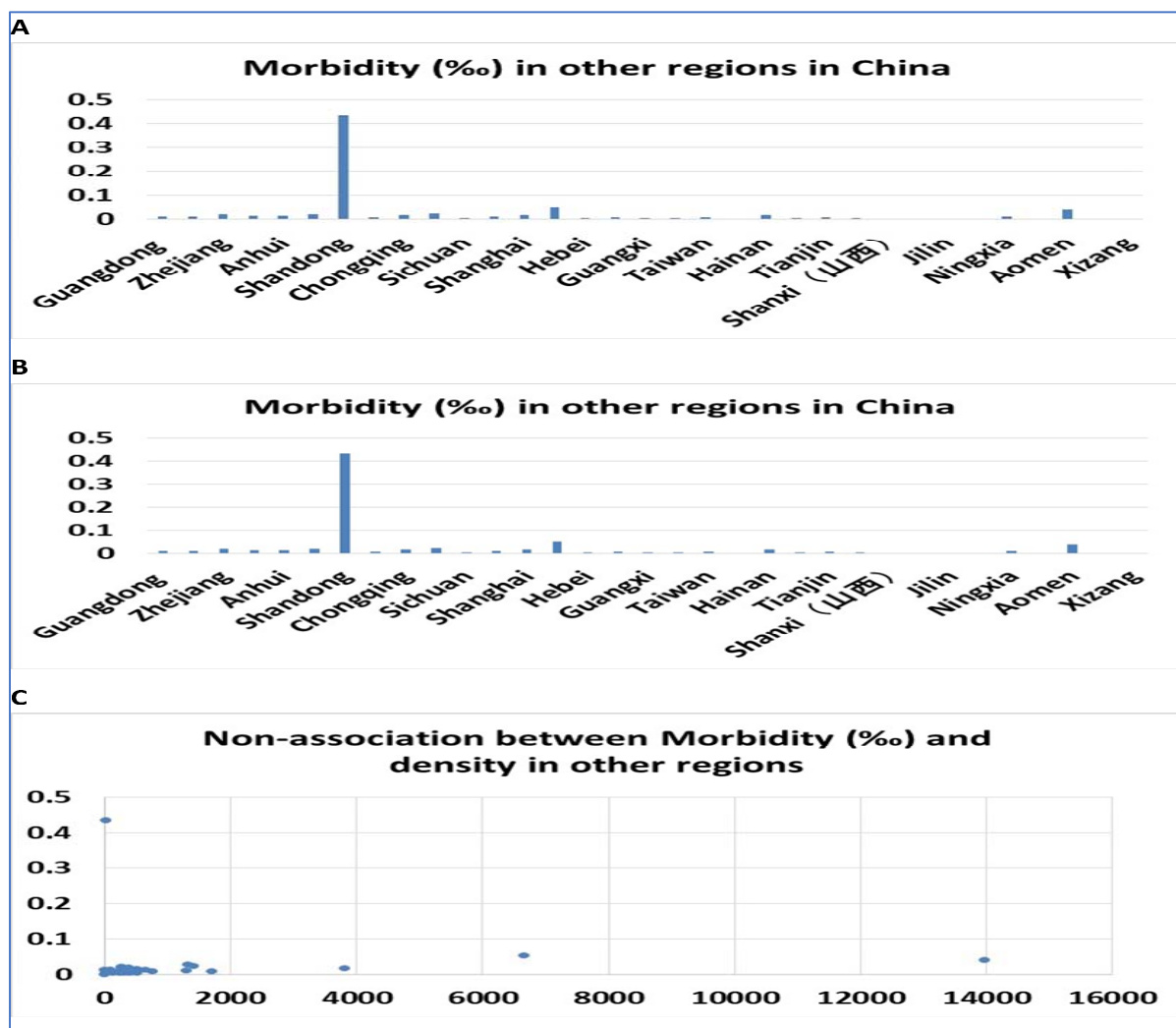
207 COVID-19. Names of the states and territories are listed on the horizontal axis. Fig. 2C. The

208 relationship between the population density and disease morbidity. The number on vertical

209 axis is the disease morbidity rate, and the horizontal axis is the number of persons per
210 kilometer.

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

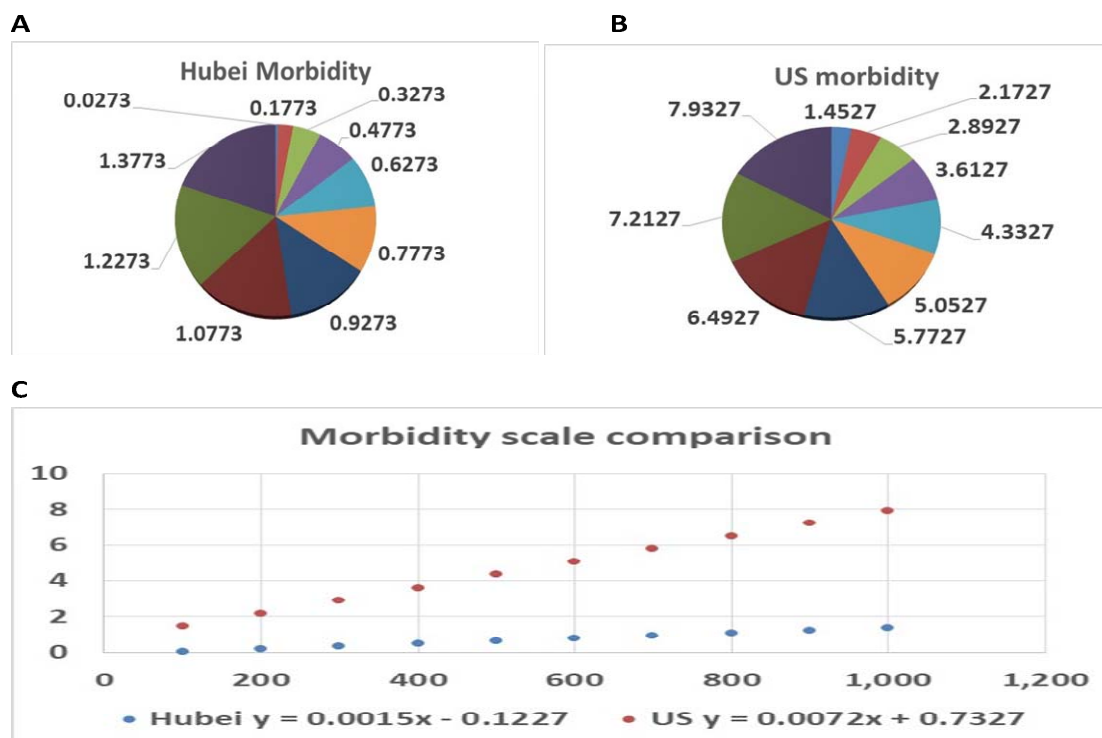
212 Since the disease in other regions of China was well under control, an epidemic did not
213 materialize into the general population before elimination of the infection source. We
214 hypothesized that in this case the population density in other regions would not be associated
215 with the disease morbidity. These regions include 33 regions, with an average of 1,106 persons
216 per kilometer but with large differences from 2.80 persons per kilometer in XiZang to 13,984
217 persons per kilometer in Aomen (Figure 3A). The disease morbidity rate is low, all below 0.1%,
218 except Shandong which had a rate of 0.4% (Figure 3B). The r value for the correlation between
219 population density and disease morbidity rate is 0.04 (Figure 3C). Thus, the data from other
220 regions in China services as a negative control; when the COVID-19 disease does not morph into
221 an epidemic, population density is not associated with the disease morbidity.



222
223 Figure 3. Non-association between population density and morbidity in other regions in
224 China. Fig. 3A: Population density in the 33 regions in China. The number on vertical axis
225 is the average number of persons per kilometer. Names of the cities are listed on the
226 horizontal axis. Fig. 3B: Disease morbidity rate in the 33 regions in China. The number on
227 the vertical axis is the average rate of morbidity. Names of the cities are listed on the
228 horizontal axis. Fig. 3C. The relationship between the population density and disease
229 morbidity. The number on vertical axis is the disease morbidity rate, and the horizontal axis
230 is the number of persons per kilometer.

231 5. Difference between China and the US on the measures of social distance and the impact of
232 disease epidemic.

233 By analyses of the data from the Hubei province in China and 53 major US cities in US states
234 and territories, we obtained the positive correlations between the population density and the
235 disease morbidity. We next asked whether there is a difference between these two data sets of
236 cities on the impact of the disease epidemic. We calculated the morbidity of population density
237 from 100 from 1000 using the linear formula obtained from the Hubei and major cities of US
238 states and territories. Although both data sets have positive correlation between the population
239 density and morbidity, they are not at the same degree. Based on the formula derived from Hubei
240 province, the morbidity increased from 0.177 to 1.377 when the population density increased
241 from the 100 to the 1000 (Figure 4A). On the other hand, in major cities and territories of US,
242 the mortality increased from 1.453 to 7.93 when the population density increased from the 100 to
243 the 1000 (Figure 4B). The increases of rates between these two sets of data are significantly
244 different (Figure 4C).



245
 246 Figure 4. Comparison of the morbidity scales between major cities in the Hubei province of
 247 China and US states and territories. Fig. 4A: The morbidity rates at ten different scales
 248 calculated based on the formula derived from 17 cities of the Hubei province in China. Fig.
 249 4B. The morbidity rates at ten different scales calculated based on the formula derived from 53
 250 major cities of US states and territories. Fig. 4C. The comparison of disease morbidity rates
 251 between cities in China and in US states and territories when the population density
 252 increases from 100 to 1000 persons per kilometer. The number on vertical axis is the
 253 disease morbidity rate, and the horizontal axis is the number of persons per kilometer.

254 **Discussion**

255 Our analyses clarified the positive correlation between population density and epidemic infection
 256 rate. From the data of 17 cities in the Hubei province, the correlation between population density
 257 and the number of people infected with COVID-19 in this epidemic is 0.62, showing a relatively

258 high positive correlation. An analysis of 53 locations in US states and territories yielded a
259 correlation of 0.55 when examining the relation between population density and the number of
260 people previously infected. Additionally, the number of infected patients is not directly related to
261 the date of initial infection. These two sets of data show that due to the higher population
262 density, the chance of contact between people is relatively high. When the coefficient
263 representing the relation between population density and number of persons infected is relatively
264 high, the absolute number of people who contact each other is relatively high. The findings from
265 this study indirectly support the practice of keeping social distance and enforcing of travel
266 restriction.

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

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

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

282

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

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302 **Acknowledgments**

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

307

308 **Author Contributions**

309 Conceived and designed the experiments: HC, WG, YJ, LA, HC, DS. Performed data searching
310 and collection: TS, HY, LY, WG. Analyzed the data: TS, HY, LY LM, LL, WG. Contributed
311 analysis tools: WG, CH, HC, DS. Wrote the manuscript: TS, HY, LY, LA, SD, CG, SH, WG.
312 Revise and approve and manuscript: All authors.

313

314 **Competing interests**

315 The authors declare no competing financial interests.

316

317

318

319

320

321

322

323

324 **References**

325

326 Bedford J, Enria D, Giesecke J, Heymann DL, Ihekweazu C, Kobinger G, Lane HC, Memish Z,
327 Oh MD, Sall AA, Schuchat A, Ungchusak K, Wieler LH; WHO Strategic and Technical
328 Advisory Group for Infectious Hazards. COVID-19: towards controlling of a pandemic. *Lancet*.
329 2020 Mar 28;395(10229):1015-1018.

330

331 Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T,
332 Zhang X, Zhang L. Epidemiological and clinical characteristics of 99 cases of 2019 novel
333 coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020 Jan 30. pii: S0140-
334 6736(20)30211-7.

335

336 Gewin V. Safely conducting essential research in the face of COVID-19. *Nature*. 2020
337 Apr;580(7804):549-550.

338

339 Gibson L, Rush D. Novel Coronavirus in Cape Town Informal Settlements: Feasibility of Using
340 Informal Dwelling Outlines to Identify High Risk Areas for COVID-19 Transmission From A
341 Social Distancing Perspective. *JMIR Public Health Surveill.* 2020 Apr 6;6(2):e18844.
342

343 Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T,
344 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
345 R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel
346 coronavirus in Wuhan, China. *Lancet.* 2020 Jan 24. pii: S0140-6736(20)30183-5.
347

348 Jarvis CI, Van Zandvoort K, Gimma A, Prem K; CMMID COVID-19 working group, Klepac P,
349 Rubin GJ, Edmunds WJ. Quantifying the impact of physical distance measures on the
350 transmission of COVID-19 in the UK. *BMC Med.* 2020 May 7;18(1):124.
351

352 Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, Tam C, Dickens BL. Interventions to
353 mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *Lancet Infect Dis.* 2020
354 Mar 23. pii: S1473-3099(20)30162-6.
355

356 Lasry A, Kidder D, Hast M, Poovey J, Sunshine G, Winglee K, Zviedrite N, Ahmed F, Ethier
357 KA; CDC Public Health Law Program; New York City Department of Health and Mental
358 Hygiene; Louisiana Department of Health; Public Health – Seattle & King County; San
359 Francisco COVID-19 Response Team; Alameda County Public Health Department; San Mateo
360 County Health Department; Marin County Division of Public Health. Timing of Community
361 Mitigation and Changes in Reported COVID-19 and Community Mobility - Four U.S.

362 Metropolitan Areas, February 26-April 1, 2020. MMWR Morb Mortal Wkly Rep. 2020 Apr
363 17;69(15):451-457.

364

365 Sun K, Chen J, Viboud C. Early epidemiological analysis of the coronavirus disease 2019
366 outbreak based on crowdsourced data: a population-level observational study. Lancet Digit
367 Health. 2020 Apr;2(4):e201-e208

368

369 Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health
370 concern. Lancet. 2020 Jan 24. pii: S0140-6736(20)30185-9.

371

372 Wang L, Li J, Guo S, Xie N, Yao L, Cao Y, Day SW, Howard SC, Graff JC, Gu T, Ji J, Gu W,
373 Sun D. Real-time estimation and prediction of mortality caused by COVID-19 with patient
374 information based algorithm. Sci Total Environ. 2020 Apr 8;727:138394.

375

