

1 **Epidemiologic Characteristics of COVID-19 in Guizhou, China**

2 **Authors:** Kaike Ping <sup>1</sup>, Mingyu Lei <sup>1</sup>, Yun Gou, Ying Tao, Yan Huang <sup>2</sup>

3 <sup>1</sup> Contributed equally to the article.

4 <sup>2</sup> Corresponding author.

5 **Affiliations:** Guizhou Center for Disease Control and Prevention, Guizhou, China (K.  
6 Ping, M. Lei, Y. Gou, Y. Tao, G. Yao, C. Hu, Z. Zou, D. Wang, Y. Huang)

7 **Address for correspondence:** Y. Huang. Institute of Communicable Disease  
8 Prevention and Control, Guizhou Center for Disease Control and Prevention,  
9 Guizhou, China; email: cdchuangyan@163.com

10

11 **Abstract**

12 At the end of 2019, a coronavirus disease 2019 (COVID-19) outbreak in  
13 Wuhan, China, and spread to Guizhou province on January of 2020. To acquire the  
14 epidemiologic characteristics of COVID-19 in Guizhou, China, we collected data on  
15 162 laboratory-confirmed cases related to COVID-19. We described the demographic  
16 characteristics of the cases and estimated the incubation period, serial interval and  
17 basic reproduction number. With an estimation of 8 days incubation period and 6 days  
18 serial interval, our results indicate that there may exist infectiousness during the  
19 incubation period for 2019-nCoV. This increases the difficulty of screening or  
20 identifying cases related to COVID-19.

21

22 **Background**

23           At the end of 2019, a Novel Coronavirus Infected Pneumonia (NCIP) outbreake  
24   in Wuhan, China. This coronavirus was named as 2019-Novel Coronavirus (2019-  
25   nCoV), and then renamed this disease caused by the virus strain from 2019-nCoV  
26   acute respiratory disease to "coronavirus disease 2019" (COVID-19) by World Health  
27   Organization (WHO) officially<sup>0</sup>. This incident caused a public health issue within  
28   China immediately, and spread to several countries over the world<sup>2-5</sup>.

29           Guizhou province is in Southwest of China, close (~1,000 km) to Wuhan. With  
30   the increasing number of COVID-19 patients in mainland China, Guizhou detected its  
31   first confirmed case on January 21, and was officially confirmed by Chinese Center  
32   for Disease Control and Prevention (China CDC) on January 23. Before the  
33   confirmation of the first case, Guizhou CDC identified and screened suspected  
34   COVID-19 cases by “pneumonia of unknown etiology” surveillance mechanism<sup>6</sup>.  
35   After the confirmation, Guizhou CDC started to identify and screen every case with  
36   pneumonia.

37           Until February 16, over 140 laboratory-confirmed cases were officially  
38   reported by Guizhou Health Committee<sup>7</sup>. In this study, we analyzed the data on the  
39   laboratory-confirmed cases in Guizhou to demonstrate the epidemiologic  
40   characteristics of COVID-19.

## 41   **Methods**

### 42   **Case Definitions**

43           According to standard clinical guidelines, suspected case of COVID-19 was  
44   defined as the combination of clinical characteristics and epidemiologic histories.  
45   Clinical characteristics of suspected case must fit at least 2 of 3 following criteria:

46 “fever and/or symptoms in respiratory system; radiographic evidence of pneumonia;  
47 low or normal white-cell count or low lymphocyte count.”<sup>8</sup> Epidemiologic histories  
48 must fit at least 1 of 4 following criteria: “A history of traveling Hubei Province or  
49 other districts that has confirmed cases reported within 14 days of symptom onset; A  
50 history of contacting with patient who has fever or symptoms in respiratory system  
51 from Hubei Province or other districts that has confirmed cases reported within 14  
52 days of symptom onset; Any person who has had a close contact with confirmed  
53 cases; cluster cases.”<sup>8</sup> With a stricter definition, Guizhou Health Committee counted  
54 all cases that fit clinical characteristics criteria but not any criteria of epidemiologic  
55 histories described above as suspected cases. A confirmed case was defined as a case  
56 with respiratory specimens that tested positive for the 2019-nCoV by at least one of  
57 the following two methods: positive result by real-time reverse-transcription–  
58 polymerase-chain-reaction (RT-PCR) assay for 2019-nCoV or a genetic sequence that  
59 matches 2019-nCoV. Asymptomatic carrier (i.e. case that displays no signs or  
60 symptoms to 2019-nCoV) was not announced by Health Committee officially, but  
61 was included in this study.

## 62 **Sources of Data**

63 Suspected case was screened by local hospital or local Center for Disease  
64 Control and Prevention (CDC). Once the patient was identified as confirmed case by  
65 laboratory, a joint field epidemiology team comprising members from Guizhou CDC  
66 and local CDC would open a detailed field investigation on demography information,  
67 epidemiologic histories, timelines of key events, and close contacts.

68 All data from epidemiological reports were inputted into standardized forms  
69 from technical protocols designed by China CDC. At least two team members

70 independently review the full report of each case report to ensure that data is correctly  
71 input.

## 72 **Epidemiological investigation**

73           Once a case was evaluated as suspected case, local CDC would finish an  
74 initial investigation report in 24 hours and collect respiratory specimens for  
75 centralized laboratory testing. When case was confirmed as positive to 2019-nCoV,  
76 province CDC would send a special epidemiology team to investigate with local  
77 CDC, in order to acquire more detailed information. Information was collected from  
78 infected individuals, family members, medical workers, close contacts, GPS-info, and  
79 CCTV camera, etc. Information included basic demography data, detailed life trace  
80 and all close contacts, clinical characteristics, exposure history, etc.

## 81 **Laboratory confirmation**

82           With the guidelines of protocol by World Health Organization (WHO)<sup>9</sup>,  
83 specimens of suspected cases were taken from whose upper and lower respiratory  
84 tract by professional medical worker from CDC or hospital. RNA was extracted and  
85 tested by RT-PCR with primers and probes for 2019-nCoV. Cross-reactivity with other  
86 known respiratory viruses and bacteria such as Influenza A (H1N1, H3N2, H5N1, or  
87 H7N9), Influenza B (Victoria or Yamagata), MERS-CoV, Adenovirus, etc. are also be  
88 tested.

## 89 **Statistical analysis**

90           The onset date was defined as the self-reported first date to have symptoms  
91 related to 2019-nCoV or the date of visiting to clinical facilities if it was unavailable  
92 to acquire the accurate onset date, and as the confirmation date for confirmed cases

93 without symptoms. As the date of infecting exposure for a given individual falls  
94 within a finite interval, an exposure window was clarified to denote this interval.  
95 Criteria of infecting exposure was described above in the **Case Definitions** section.  
96 Continuous variables were expressed as the means and standard deviations or  
97 medians, the interquartile range (IQR) equals the difference between 75th and 25th  
98 percentiles. Categorical variables were summarized as the counts and percentages in  
99 each category. The incubation period is defined as the delay from viral infection to the  
100 onset of illness<sup>10</sup>, which was estimated by fitting a parametric accelerated failure time  
101 model with log-normal distribution of cases with detailed exposure window data, and  
102 standard errors of parameters was computed using a 1,000 times bootstrap routine,  
103 this was performed by R package *coarseDataTools*<sup>11</sup>. The serial interval distribution  
104 (i.e. the duration between symptom onset of a primary/index case and symptom onset  
105 of its secondary cases) was fitted by a Weibull distribution with data from cluster  
106 events. The basic reproduction number ( $R_0$ ) was defined as the expected number of  
107 cases directly generated by one case in a population where all individuals are  
108 susceptible to infection, it was estimated by maximum likelihood estimation via R  
109 package *R0*<sup>12</sup>.

110 All analyses and statistical graphs were conducted with R software version  
111 3.6.2 (R Foundation for Statistical Computing).

## 112 **Ethics approval**

113 Data collection and analysis of cases and their close contacts were determined  
114 by Guizhou Health Commission of the People's Republic of China. It is part of a  
115 continuing public health outbreak investigation and exempted from institutional  
116 review board assessment.

## 117 **Results**

118           There were 759 cases in total. Among all patients, 144 (18.97%) cases were  
119 confirmed cases with symptoms to 2019-nCoV, 18 (2.37% of all 759 cases, 11.11%  
120 of all 162 confirmed cases) cases were confirmed as cases without symptoms, 597  
121 (78.65%) cases were diagnosed as other diseases after clinical diagnosing or  
122 laboratory testing, 34 (4.47%) cases were suspected cases that were pending for  
123 laboratory testing. **Figure 1** shows the geographic distribution of confirmed cases  
124 throughout China and Guizhou. The median age of all confirmed cases was 37 years  
125 (ranges from 1 month to 89 years), 82 (50.62%) confirmed cases were male, 84  
126 (51.85%) confirmed cases were exposed to Wuhan or other districts outside Guizhou,  
127 the detailed demographic data are shown in **Table 1**.

128           The first case was detected on January 21 and was officially confirmed by  
129 China CDC on January 23. It was reported by Guiyang (capital city of Guizhou)  
130 CDC. This case was a 51-year-old male, infected in Wuhan where he was there to do  
131 some business affairs. His onset date was on January 9 when he was still in Wuhan.  
132 After returning to Guizhou on January 14, he visited a local hospital for several times  
133 until January 16 when he was isolated by the Division of Infection. **Figure 2** shows  
134 the distribution of the onset date of all cases related to 2019-nCoV over time. It is  
135 notable that there was a decline of cases after February 7th, which is most likely  
136 attributed to the delay of reporting by the local hospital and laboratory confirmation.

137           We included 93 confirmed cases with detailed exposure window and onset  
138 date information to estimate the incubation period (Fig. 3 A-B). The median  
139 incubation period was 8.06 days (95% confidence interval [CI]: 6.89 – 9.36). The  
140 95% percentile was 21.90 days (95% CI: 18.21 – 25.45). Other percentiles are shown

141 on **Table 2**. We obtained 57 observations from cluster cases and estimated the serial  
142 interval distribution (Fig. 3C) with the mean ( $\pm$ SD) of  $6.37\pm 4.15$  days. The estimated  
143 basic reproduction number ( $R_0$ ) was 1.09 (95% CI: 0.94 – 1.26).

## 144 **Discussions**

145 This study provides an initial analysis among epidemiologic characteristics  
146 and two typical transmission phenomena of COVID-19. Most index cases were  
147 exposed to Wuhan or confirmed cases related to Wuhan, but it is also necessary to  
148 prevent the community transmission of COVID-19. Measures such as community  
149 screening, isolation of close contacts, public transportation limitation, extended  
150 population of PCR testing, enhanced disinfection, and health education etc. could be  
151 taken. Our findings provide more epidemiologic data to this incident of COVID-19,  
152 which contribute to the further analysis and the control of this disease.

153 In addition to the evidence of the human-to-human transmission, it is likely  
154 that there exists infectiousness to some extent during the incubation period, and the  
155 presence of “super-spreaders”. The median incubation period was 8.2 days (95% CI:  
156 7.9 – 9.5) in our study, which is longer than a recent report of 425 patients (8.2 days  
157 vs. 5.2 days), this may be a results of recall bias, during epidemiological investigation,  
158 we found that some cases actually have had shown mild symptoms days before the  
159 date they reported so as to their ignorance. We estimated the  $R_0$  as 1.09 (95% CI: 1.9  
160 – 1.2), which means each patient can infect other 1.09 people on average. In  
161 commonly used infection models, when  $R_0 > 1$  the infection will be able to start  
162 spreading in a population, but not if  $R_0 < 1$ .

163 However, it should be treated with caution that the basic reproduction number  
164  $R_0$  was calculated under the assumptions that an infected individual but has no

165 symptoms does not yet infect others, which assumes that all index cases should show  
166 symptoms before their secondary cases. However, as the investigation went on,  
167 Guizhou CDC found that there existed index cases who showed symptoms after that  
168 of their secondary cases. This may lead to that the estimated  $R_0$  was not robust as we  
169 do not hold enough evidence and information to the presence of asymptomatic carrier,  
170 those data were not included in the estimation model.

171 There are several limitations in our study. First, we could not collect all  
172 complete exposure and onset data from all confirmed cases, due to the recall bias and  
173 the lack of standardization on epidemiological investigation at the beginning of the  
174 incident. Second, some key data from asymptomatic carrier, especially the onset date  
175 or onset window, and left bound of exposure window from cases of Wuhan residents  
176 were unavailable, other statistical model may be used to fit such kind of censored data  
177 to estimate serial interval and  $R_0$  more accurately.

## 178 **Funding**

179 The authors received no specific funding for this research.

## 180 **Author Bio**

181 Kaike Ping acquires his Master degree of Biostatistics in Southern Medical  
182 University on 2017. Now he is a researcher at Center for Disease Control and  
183 Prevention of Guizhou, China, where he has been investigating respiratory diseases  
184 and their epidemiologic Characteristics. His research interests include disease  
185 distributions, time series analysis, survival analysis, and machine learning.

## 186 **References**

187 1. World Health Organization. Novel Coronavirus (2019-nCoV): situation report,

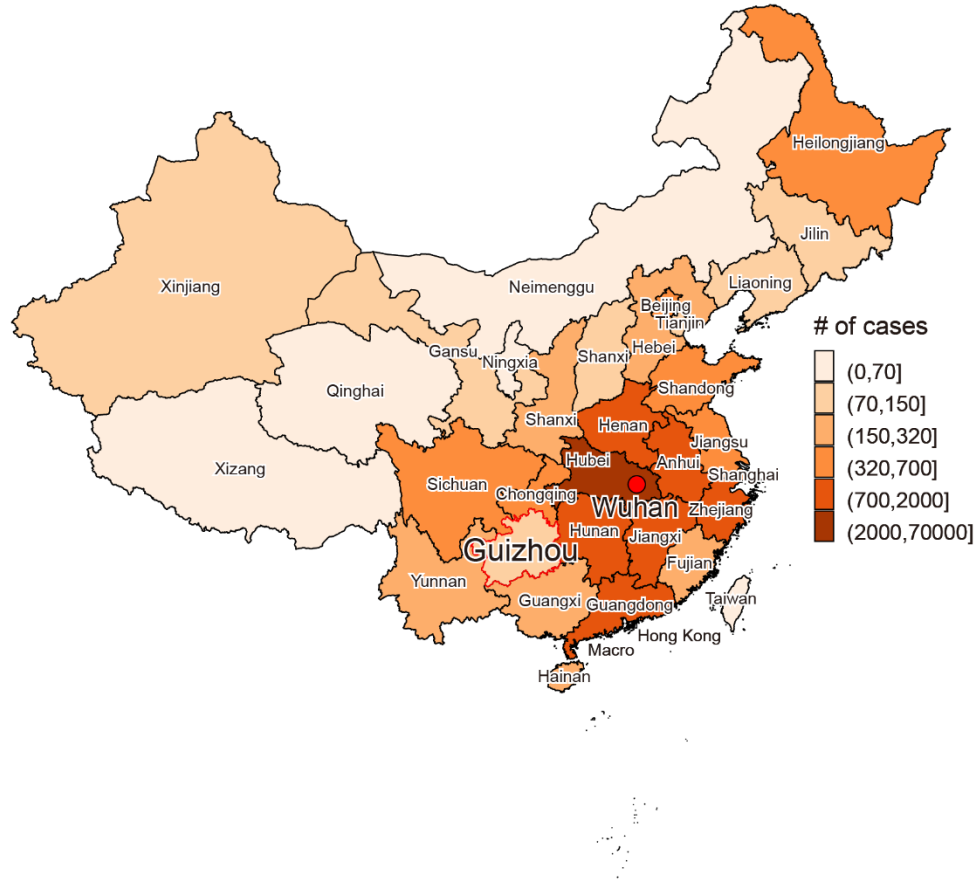


- 188 22 [EB/OL]. World Health Organization. [https://www.who.int/docs/default-](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf)  
189 [source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf). 2020-02-  
190 11.
- 191 2. Ryan R, Kerri B. Health officials to announce 'presumptive case' of  
192 coronavirus in Toronto [EB/OL]. Global News.  
193 <https://globalnews.ca/news/6462626/coronavirus-toronto-hospital/>. 2020-01-  
194 25.
- 195 3. Lena H. S, Lenny B. First U.S. case of potentially deadly Chinese coronavirus  
196 confirmed in Washington state [EB/OL]. The Washington Post.  
197 <https://www.washingtonpost.com/health/2020/01/21/coronavirus-us-case/>.  
198 2020-01-21.
- 199 4. World Health Organization. WHO | Novel Coronavirus – Japan (ex-China)  
200 [EB/OL]. World Health Organization. [http://www.who.int/csr/don/16-january-](http://www.who.int/csr/don/16-january-2020-novel-coronavirus-japan-ex-china/en/)  
201 [2020-novel-coronavirus-japan-ex-china/en/](http://www.who.int/csr/don/16-january-2020-novel-coronavirus-japan-ex-china/en/). 2020-01-17.
- 202 5. Jacob E. Coronavirus: trois premiers cas confirmés en France [EB/OL]. Le  
203 Figaro.fr (in French). [https://www.lefigaro.fr/sciences/coronavirus-trois-](https://www.lefigaro.fr/sciences/coronavirus-trois-premiers-cas-confirmes-en-france-20200124)  
204 [premiers-cas-confirmes-en-france-20200124](https://www.lefigaro.fr/sciences/coronavirus-trois-premiers-cas-confirmes-en-france-20200124). 2020-01-24.
- 205 6. Xiang N, Havers F, Chen T, et al. Use of national pneumonia surveillance to  
206 describe influenza A(H7N9) virus epidemiology, China, 2004–2013. *Emerg*  
207 *Infect Dis* 2013; 19: 1784-90.
- 208 7. Guizhou Health Committee. COVID-19 condition in Guizhou province until  
209 2020-02-14 [EB/OL]. Guizhou Health Committee.  
210 [http://www.gzhfpc.gov.cn/xwzx\\_500663/zwyw/202002/t20200215\\_49183203](http://www.gzhfpc.gov.cn/xwzx_500663/zwyw/202002/t20200215_49183203).

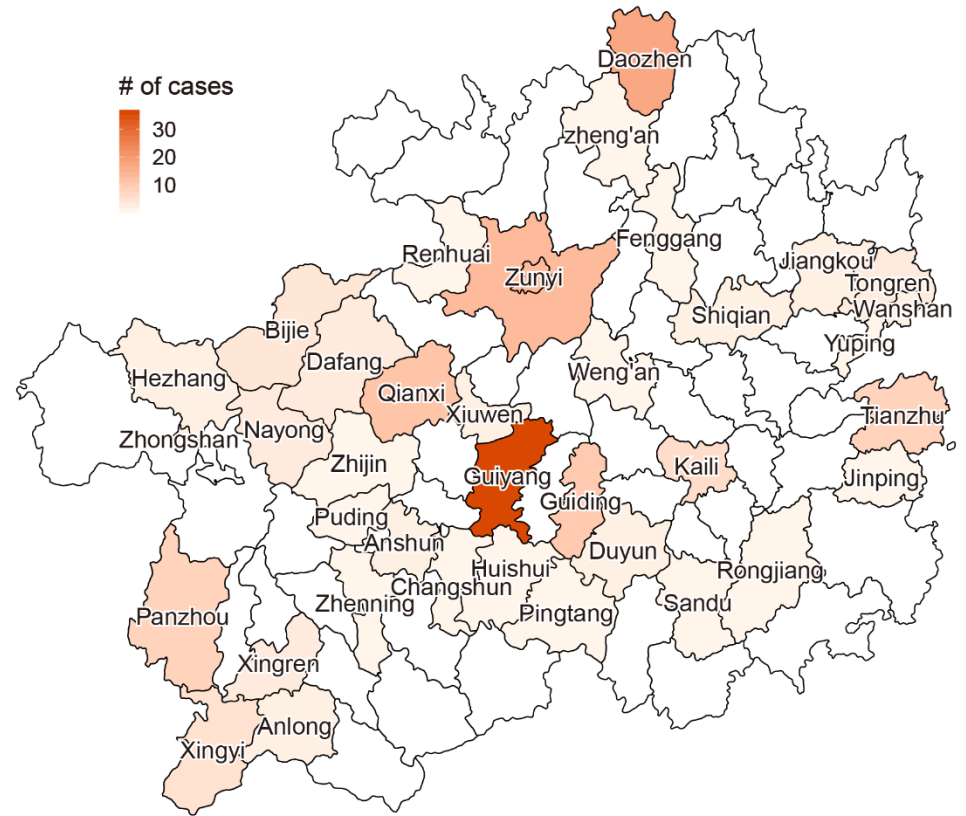
- 211           html. 2020-02-15.
- 212           8. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of  
213           Novel Coronavirus–Infected Pneumonia[J]. *New England Journal of*  
214           *Medicine*, 2020.
- 215           9. World Health Organization. Laboratory diagnostics for novel coronavirus  
216           [EB/OL]. World Health Organization. [https://www.who.int/health-](https://www.who.int/health-topics/coronavirus/laboratory-diagnostics-for-novel-coronavirus)  
217           [topics/coronavirus/laboratory-diagnostics-for-novel-coronavirus](https://www.who.int/health-topics/coronavirus/laboratory-diagnostics-for-novel-coronavirus). 2020-02-12.
- 218           10. Vynnycky E, Fine PE. Lifetime risks, incubation period, and serial interval of  
219           tuberculosis. *Am J Epidemiol*. 2000;152(3): 247–263.
- 220           11. Reich NG, Lessler J, Cummings DAT, Brookmeyer R. (2009). Estimating  
221           incubation periods with coarse data. *Statistics in Medicine*. 28(22):2769--  
222           2784.
- 223           12. Obadia T, Haneef R, Boëlle P Y. The R0 package: a toolbox to estimate  
224           reproduction numbers for epidemic outbreaks. *BMC medical informatics and*  
225           *decision making*, 2012, 12(1): 147.
- 226           13. Rothe, Camilla, et al. "Transmission of 2019-nCoV infection from an  
227           asymptomatic contact in Germany." *New England Journal of Medicine* (2020).
- 228           14. National Health Committee. COVID-19 condition in China until 2020-02-15  
229           [EB/OL]. National Health Committee [EB/OL].  
230           <http://www.nhc.gov.cn/yjb/s7860/202002/4a1b1ec6c03548099de1c3aa935d04>  
231           fd.shtml. 2020-02-16.

- 232 15. Nishiura, Hiroshi, et al. "The Rate of Underascertainment of Novel  
233 Coronavirus (2019-nCoV) Infection: Estimation Using Japanese Passengers  
234 Data on Evacuation Flights." (2020): 419.  
235

A



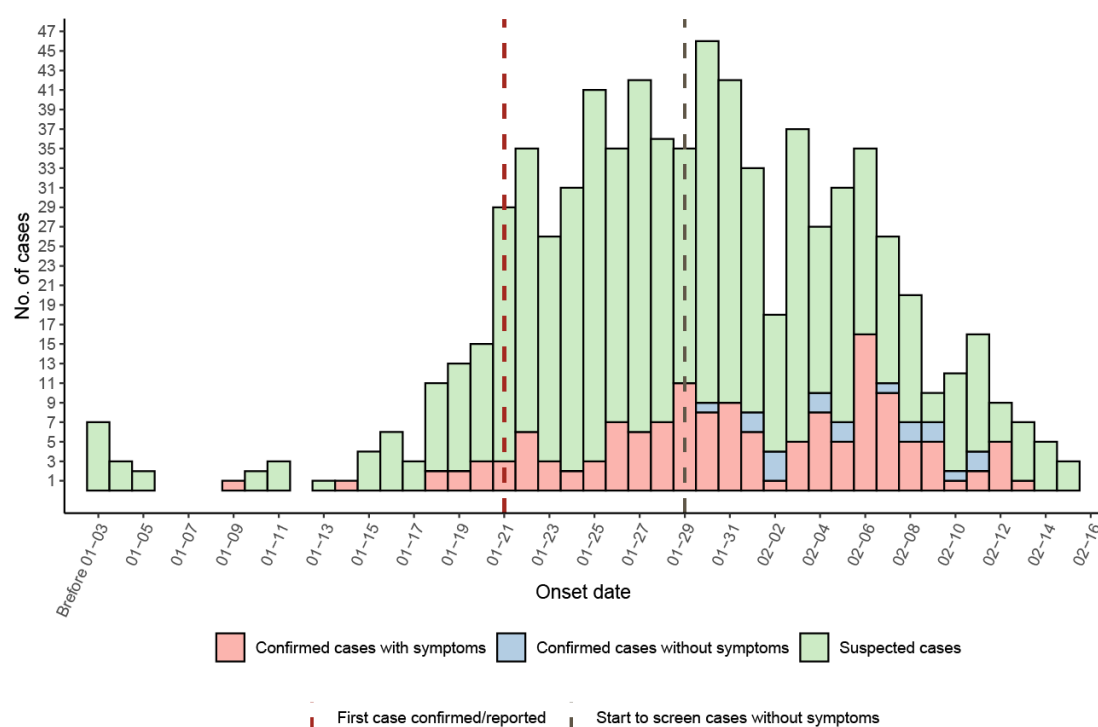
B



236 **Figure 1. The geographic distribution of confirmed cases in China and Guizhou.**

237 **(A) Geographic distribution** and statistics of all confirmed cases in China as of February 16. Data was collected from National Health

238 Committee<sup>14</sup>. In total 57,416 cases were confirmed throughout China, and 144 cases  
 239 were in Guizhou. **(B) Geographic distribution and** statistics in Guizhou province as  
 240 of February 16. In addition to confirmed cases with symptom, 18 cases without  
 241 symptom were also included, leading to 162 cases shown in the figure. Guiyang,  
 242 locating in the center of Guizhou, is the capital and largest city of the province, where  
 243 38 cases were reported. Throughout Guizhou province, 38 out of 88 administrative  
 244 districts have reported confirm cases.

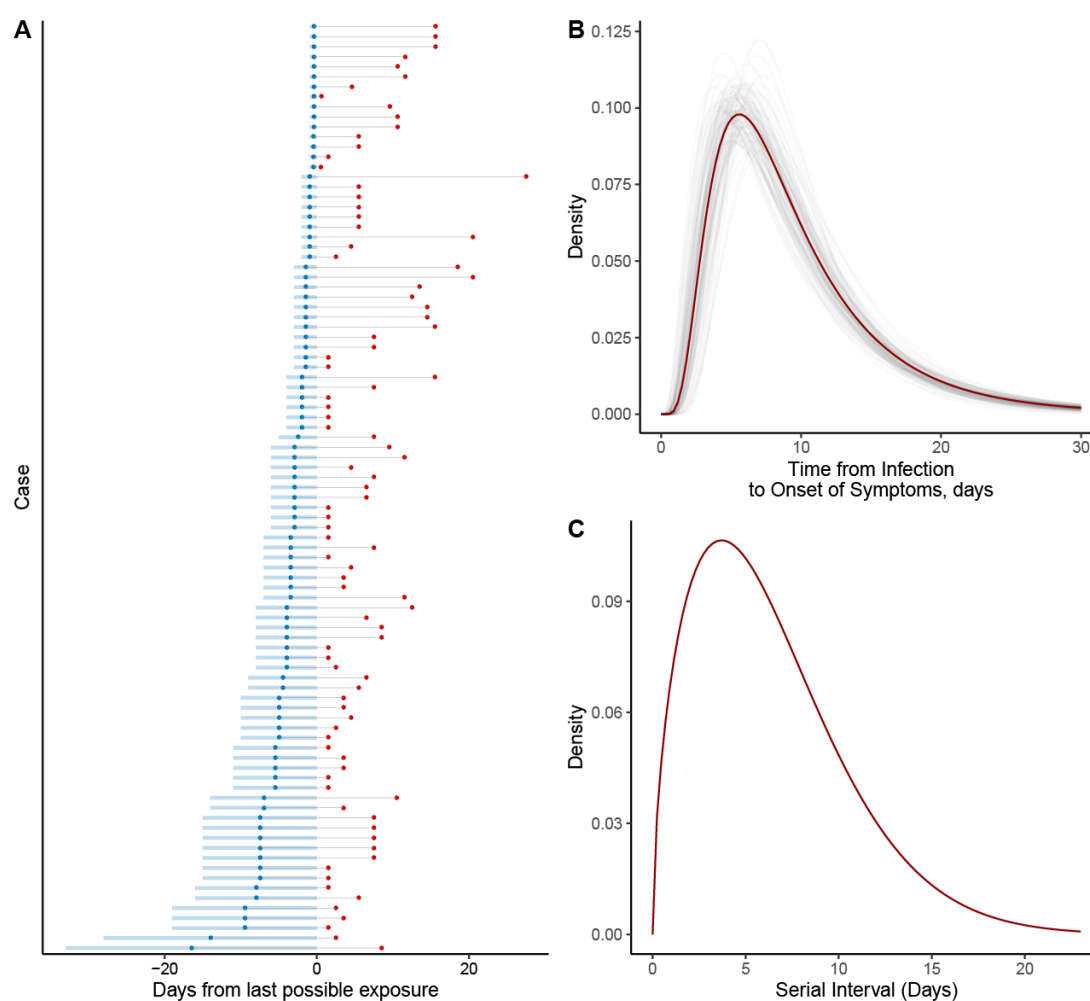


245

246 **Figure 2. Distribution of the onset date of confirmed and suspicious cases of**  
 247 **COVID-19 in Guizhou, China.**

248 The horizontal axis represents the onset date of confirmed cases or suspected cases.  
 249 The onset data was defined as the self-reported first date to have symptoms related to  
 250 2019-nCoV and as the confirmation date for confirmed cases without symptoms. On  
 251 January 21, the first case was officially confirmed and reported by Guiyang (capital  
 252 city of Guizhou) CDC, whose onset date was on January 9th when the case was still in

253 Wuhan. On January 29th, we started to identify and screen close contacts of  
254 confirmed cases and those had a positive result by real-time reverse-transcription–  
255 polymerase-chain-reaction (RT-PCR) yet without symptoms were labeled as  
256 “Confirmed cases without symptoms” in the figure. It is notable that there shows a  
257 decline of cases after February 12, which is most likely attributed to the delay of  
258 reporting by the local hospital and laboratory confirmation.



260 **Figure 3. Key time-to-event distributions.**

261 **(A)** The time windows of exposure and onset points of symptoms for 93 cases (45  
262 females, 48.38%; 48 males, 51.62%; median age 42, IQR: 27-55) that have detailed  
263 information of exposure and onset dates. Blue shaded regions indicate the exposure

264 windows, and blue points represent the midpoint within the region. Exposures were  
 265 defined as travel histories to Wuhan or close contact to other infectious individuals.  
 266 **(B)** The estimated incubation period distribution using a log-normal model with 1,000  
 267 times bootstrap resamples. The estimated median incubation period in Guizhou is 8.06  
 268 days (95% CI: 6.89 – 9.36). It was estimated that approximately 2.5% of the  
 269 individuals have symptoms in 2.45 days (95% CI: 1.71 – 3.56) after the infection, and  
 270 97.5% of that show symptoms within 26.52 days (95% CI: 21.35 – 31.77) after the  
 271 infection. **(C)** The estimated serial interval distribution -- the duration between  
 272 symptom onset of a primary case and symptom onset of its secondary cases -- fitted  
 273 via the Weibull distribution, with a mean ( $\pm$ SD) of 6.37 $\pm$ 4.15 days.

274 **Table 1.** Demographic Characteristics of Cases with COVID-19 in Guizhou

Characteristics	Value
Sample size	162
Cases with symptoms	144
Cases without symptoms	18
Median age (range) — yr	37 (1 month - 89)
Age group — no. (%)	
<15 yr,	16 (9.88%)
15–44 yr,	88 (54.33%)
45–64 yr,	46 (28.39%)
$\geq$ 65 yr	12 (7.40%)
sex	
Male	82 (50.62%)
Female	80 (49.38%)
Exposure info — no. (%)	
Detailed exposure window	93 (58.02%)
Only right bound of exposure window known	54 (33.33%)
exposure window unknown	4 (2.47%)
Onset date unknow	11 (6.18%)
Exposure in Wuhan or other districts outside Guizhou— no. (%)	84 (51.85%)

275 **Table 2.** Percentiles of incubation period

	Estimates	95%CI Low	95%CI High	Standard Error
P2.5	2.455	1.716	3.565	0.485

P5	2.972	2.15	4.162	0.523
P25	5.358	4.302	6.625	0.611
P50	8.069	6.897	9.362	0.636
P75	12.152	10.64	13.514	0.733
P95	21.904	18.219	25.452	1.833
P97.5	26.523	21.356	31.775	2.635

---

276 -2\*Log Likelihood = 306.3