

1 Title: **Intrinsic growth rules of patients infected, dead and cured with 2019 novel**  
2 **coronavirus in mainland China**

3 Chuanliang Han<sup>1,Δ</sup>, Yimeng Liu<sup>2,3,4,Δ</sup>, Saini Yang<sup>2,3,4\*</sup>

4  
5 <sup>1</sup>State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern  
6 Institute for Brain Research, Beijing Normal University, Beijing 100875, China

7 <sup>2</sup>Key Laboratory of Environmental Change and Natural Disaster, Ministry of  
8 Education, Beijing Normal University, Beijing 100875, China

9 <sup>3</sup>State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing  
10 Normal University, Beijing 100875, China

11 <sup>4</sup>Academy of Disaster Reduction and Emergency Management Ministry of  
12 Emergency Management and Ministry of Education, Faculty of Geographical Science,  
13 Beijing Normal University, Beijing 100875, China

14  
15 <sup>Δ</sup>C.L.H and Y.M.L contributed equally to this work

16  
17 **\*Correspondence authors:**

18 Saini Yang,  
19 Key Laboratory of Environmental Change and Natural Disaster, Ministry of  
20 Education & State Key Laboratory of Earth Surface Processes and Resource Ecology  
21 & Academy of Disaster Reduction and Emergency Management Ministry of  
22 Emergency Management and Ministry of Education, Faculty of Geographical Science,  
23 Beijing Normal University, Beijing 100875, China

24 Email: [yangsaini@bnu.edu.cn](mailto:yangsaini@bnu.edu.cn)

25  
26 **Conflict of Interest:** The authors declare no competing conflicts of interest.

27 **Acknowledgements:** This study is sponsored by the National Key Research and  
28 Development Program of China (2018YFC1508903), the National Natural Science  
29 Foundation of China (41621061) and the support of International Center for

30 Collaborative Research on Disaster Risk Reduction (ICCR-DRR).

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

32 **Abstract**

33 **Background**

34 An outbreak of a novel coronavirus (SARS-CoV-2)-infected pneumonia (COVID-19)  
35 was first occurred in Wuhan, China, in December 2019 and then spread rapidly to  
36 other regions. It has been declared that at least one confirmed case infected by  
37 SARS-CoV-2 was found in each province of China by late January 2020.

38 **Methods**

39 We collected the time series data of the cumulative number of confirmed infected,  
40 dead, and cured cases from the health commissions in 31 provinces in mainland China.  
41 A simple descriptive model in a logistic form was formulated to infer the intrinsic  
42 epidemic rules of COVID-19. Furthermore, we compared the intrinsic epidemic rules  
43 of COVID-19 in Hubei with that of severe acute respiratory syndrome (SARS) in  
44 Beijing, which was obtained from the Ministry of Public Health of China in 2003.

45 **Results**

46 The feature of data of three population groups in all the provinces could be well  
47 captured ( $R^2 > 0.95$ ) by the proposed descriptive model with slight variation among  
48 different provinces. By comparing fitted parameters, we found that the fitted curve of  
49 infected case is the earliest to be saturated and has the lowest semi-saturation period  
50 in both COVID-19 and SARS. While all three population groups of SARS is later to  
51 saturate and has longer semi-saturation period than that of COVID-19.

52 **Conclusions**

53 Our model is robust and stable to depict the intrinsic growth rule for the cumulative  
54 number of confirmed infected, dead, and cured cases in 30 provinces, which could  
55 depict the epidemic situation in mainland China very well. Despite the virus caused  
56 SARS (SARS-CoV) and the virus caused COVID-19 (SARS-CoV-2) are homologous,  
57 their epidemiologic characteristics are quite different. The duration of the outbreak  
58 would be shorter for COVID-19.

59

## 60 **Introduction**

61 In December 2019, a novel coronavirus-infected pneumonia causing human infection  
62 was first identified in Wuhan, a city of Hubei province located in central China with a  
63 population of 11 million people (Zhu et al., 2020; Hui et al., 2020). The coronavirus  
64 was firstly named as 2019-nCoV on January 12, 2020 and may have a probable bat  
65 origin (Zhou et al., 2020). Human to human transmission through 2019-nCoV has  
66 been confirmed among closed contacts since the middle of December 2019 (Wang et  
67 al. 2020; Li et al., 2020). As of February 12, 2020, infections of 2019-nCoV have been  
68 identified in all Chinese provincial administrative regions and in dozens of countries  
69 around the world (Munster et al., 2020; Holshue et al, 2020; Ki et al., 2020;  
70 Shigemura et al., 2020; Nshiera et al., 2020). So far, the coronavirus has been  
71 renamed as COVID-19 on February 12, 2020 officially. Different from a previous  
72 famous beta coronavirus in 2003, severe acute respiratory syndrome coronavirus  
73 (SARS-CoV) (Ksiazek et al., 2003; Kuiken et al., 2003; Drosten et al., 2003) that has  
74 a mortality rates of 10% (WHO), that of COVID-19 is 3.25% in Hubei province, and  
75 0.71% in rest of mainland China (from Chinese Center for Disease Control and  
76 Prevention, CCDC) up to February 19, 2020.

77

78 Here, we used a descriptive model to fit data on the 2019-nCoV cases (infection, death,  
79 cure) from January 20, 2020 to February 19, 2020, aiming to capture the  
80 epidemiologic characteristics and dynamics of NCIP in 31 provinces in Mainland  
81 China. Furthermore, we compared these characteristics with SARS in 2003 using the  
82 same model structure in order to validate this model and understand the differences  
83 among these two diseases.

84

## 85 **Method**

### 86 **Sources of data**

87 We obtained the time series data of 2019-nCoV cases from the National Health  
88 Commission of China, and the provincial Health Commission of 31 provincial  
89 administrative regions in Mainland China (January 20 to February 19, 2020). The data  
90 used in this study include the number of confirmed 2019-nCoV cases, the cumulative  
91 number of suspected, dead, and cured cases. All cases were laboratory confirmed  
92 following the case definition by national health commission of China (from CCDC  
93 website). The basic test procedure has been described in detail in previous work  
94 (Huang et al., 2020; Li et al., 2020). We only chose the cases that are tested positive  
95 by real-time RT-PCR with specific primers and probes to SARS-CoV-2 both at the  
96 corresponding provincial CDC and the National Institute for Viral Disease Control at  
97 CCDC.

98

99 The severe acute respiratory syndrome (SARS) in 2003 was another fatal coronavirus  
100 epidemics over the last two decades, we collected the data (the cumulative number of  
101 suspected, dead, and cured cases) of SARS during March 1 to August 16, 2003, in  
102 Beijing from the National Health and Family Planning Commission of PRC.

103

### 104 **Epidemic curve modeling**

105 We assumed that the population of each region was susceptible to COVID-19, the  
106 number of susceptible people per unit time was proportional to the number of  
107 uninfected persons, and isolation or closure of the city could affect the spread effect of  
108 infected to susceptible people to a certain extent. Previous studies have showed that  
109 daily cumulative number of patients could be explained by logistic function (Huang et  
110 al., 2003; Wang et al., 2003). Based on the above assumption and previous studies, we  
111 modeled the epidemic information of virus as logistic form (Eq. (1))

112 
$$\begin{cases} \frac{dI(t)}{dt} = \alpha I(t) \left(1 - \frac{I(t)}{N_1}\right) \\ I(0) = I_0 \end{cases} \quad (1)$$

113 where  $N_1$  is maximum number of cumulative infections,  $I(t)$  is cumulative number of  
114 patients at time  $t$ ,  $\alpha$  is the incidence growth rate, and  $I_0$  is the number of infected  
115 cases at the initial time.

116 Analytical solution of Eq. (1) could be written in the form of Eq. (2)

117 
$$I(t) = \frac{N_1}{1 + C_1 e^{-\alpha N_1 t}} \quad (2)$$

118 
$$C_1 = \left(\frac{N_1}{I_0} - 1\right) e^{\alpha N_1 t_0} \quad (3)$$

119 Similarly, we could write the analytical solutions for the dead and cured cases by Eq.  
120 (4) and Eq. (6).

121 
$$D(t) = \frac{N_2}{1 + C_2 e^{-\beta N_2 t}} \quad (4)$$

122 
$$C_2 = \left(\frac{N_2}{D_0} - 1\right) e^{\beta N_2 t_0} \quad (5)$$

123 
$$R(t) = \frac{N_3}{1 + C_3 e^{-\gamma N_3 t}} \quad (6)$$

124 
$$C_3 = \left(\frac{N_3}{R_0} - 1\right) e^{\gamma N_3 t_0} \quad (7)$$

125 Where  $\beta$ ,  $\gamma$  is the growth rate of dead cases and the growth rate of cured cases,  
126 respectively.  $N_2$  and  $N_3$  is maximum number of cumulative dead and cured cases.  $D_0$   
127 and  $R_0$  is the number of dead and cured cases at initial time.

128 Here we simplified Eqs. (2),(4),(6) and unified them as Eq. (8)

129 
$$N(t) = \frac{A}{1 + e^{-k(t-t_0)}} \quad (8)$$

130 where  $N(t)$  is the general form of  $I(t)$ ,  $D(t)$  or  $R(t)$ ,  $A$  denotes  $N_1$ ,  $N_2$ , or  $N_3$ , and  $k$   
131 represents  $\alpha$ ,  $\beta$ , or  $\gamma$ , indicating the intrinsic rule for growth of cases.  $t_0$  represents  
132 the semi-saturation period, illustrating the inflection point of the sigmoid curve.

133

134 The data processing and modeling were performed on MATLAB (The MathWorks)  
135 with custom scripts. The nonlinear least square (NLS) algorithm was adopted for data  
136 fitting and parameter estimation. We used the MATLAB function “nlinfit” to

137 minimize the sum of squared differences between the data points and the fitted values.

138 Eq. (8) was used in this work as the descriptive model to depict the intrinsic growth

139 rule for infected, dead and cured cases.

140

141 **Results**

142 As of Feb 19, 2020, laboratory-conformed COVID-19 infections have been reported  
143 in each of the 31 Chinese provincial administrative region. So far, there have been a  
144 total of 75002 confirmed cases of 2019-nCoV infections in mainland China, including  
145 2118 deaths and 16157 cases cured (From CCDC). Hubei province contributed a  
146 majority of cases to this statistical numbers (Fig 1A). Based on the limited infected  
147 case data in the early and middle phase of this new coronavirus outbreak, we firstly  
148 established a descriptive model with three parameters for each of the 31 provinces in  
149 mainland China by (Fig 1B & Fig 2B, See Methods). Fitting goodness of models was  
150 expressed by the R-square, which ranged from 0.95 to 0.999, indicating that the  
151 models we established fitted the data well. The panel B of Figure 1 illustrates  
152 examples of time series data of infected cases and corresponding fitting curves in four  
153 provinces (provinces of Xinjiang, Beijing, Guangdong, Zhejiang and Heilongjiang). It  
154 is clear that different province varied in patterns of increase of infected cases.

155

156 ----- Insert Figure 1 near here -----

157

158 In our model,  $k$  is an important parameter to depict the intrinsic growth rule of  
159 cumulative number of infected patients (Eq. (8)). Based on our assumption, the  
160 growth pattern was supposed to follow a sigmoid form, and the parameter  $k$  here  
161 denotes how rapid the number of infected cases get to saturate. If  $k$  is large, the  $S$  form  
162 of the curve would be sharp, and the curve will rise up very quickly to be saturated to  
163 the maximum number. While if  $k$  is only slightly larger than 0, the curve would show  
164 a slow and smooth increase from 0 to the maximum.

165

166 To further illustrate the spatial distribution of the infected cases' growth pattern, we  
167 exerted further analyses of the fitted infected case curves of 30 provinces. We  
168 excluded Tibet in this analysis because the number of confirmed infected case did not

169 increase since Jan 29, 2020. And the only one confirmed case in Tibet has been cured  
170 on Feb 12. The spatial distribution of parameter  $k$  is shown in Fig 2A, illustrating  
171 provinces varied in the increase patterns of cumulative infected cases. Fitted curves of  
172 provinces such as Jilin and Qinghai has larger value of  $k$ , which means that the  
173 number of infections peaked earlier in these regions than in other provinces. By  
174 contrast, provinces like Hainan, Hebei, Ningxia, has smaller value of  $k$  than other  
175 provinces, suggesting the infections of these provinces would continue increasing  
176 when infections of other provinces reach saturation (Fig 2B). Fig 2C shows the spatial  
177 distribution of predicted maximum number of infected cases (parameter  $A$ ) in each  
178 province of mainland China. Our model shows that the maximum number of infected  
179 COVID-19 cases in mainland China would be 81498 (95% confidence interval:  
180 [79099, 83896]). Among all the provinces in mainland China, the Hubei province is  
181 predicted to have largest number of infected COVID-19 cases, which would be 68056  
182 (95% confidence interval: [65679,70432]). Despite Tibet is predicted to have only one  
183 infected case, the model predicted Qinghai province would have 19 cases (95%  
184 confidence interval: [18, 19]), which is the fewest compared with other provinces in  
185 mainland China.

186

187 ----- Insert Figure 2 near here -----

188

189 Moreover, to compare the intrinsic growth rule of infected cases, dead cases and cured  
190 cases between COVID-19 and SARS, we applied the descriptive model to fit the data  
191 of SARS in Beijing 2003 (Fig 3A). Since the time series data of SARS was full phase  
192 of the outbreak, the parameters estimated by our model was solid (R-square>0.95).  
193 We found the  $k$  value of fitted curves of infected SARS cases in Beijing was 0.18,  
194 which was much lower than that of Hubei province of COVID-19 ( $k = 0.25$ ) (Fig 3  
195 ABC). Furthermore, we found that all 30 provinces'  $k$  value were significantly larger  
196 than 0.18 (Fig 3D). We also noted the  $t_0$  of fitted curves of infected SARS cases in



197 Beijing was larger than that of fitted curves of infected COVID-19 cases in each  
198 province in mainland China (Fig 3E). Moreover, the increasing rate ( $k$ ) was negatively  
199 correlated (Pearson,  $r=-0.47$ ,  $p=0.0072$ ) to semi-saturation period (Fig 3F). We  
200 concluded that at the provincial level, the COVID-19 spread more quickly than the  
201 SARS did. The infections of COVID-19 increased more rapid and would reach  
202 saturation earlier than SARS according to our model.

203

204 ----- Insert Figure 3 near here -----

205

206 The outbreak of SARS in Beijing was the most serious among all the provincial  
207 administrative regions, and the outbreak of COVID-19 was most serious in Hubei  
208 province. Here we used our descriptive model to fit the dead and cured cases' curves  
209 for SARS in Beijing and COVID-19 in Hubei province (Fig 4 A-C, F-H, R-square $\geq$   
210 0.95), and compared the growth features of three types cases of SARS and COVID-19  
211 by two parameters,  $k$  and  $t_0$ (Fig 4 DEIJ). We found that  $k$  values of the three types  
212 cases of SARS were lower than that of COVID-19 (Fig 4 DE), and semi-saturation  
213 period for them in SARS was higher than COVID-19's, which suggested the duration  
214 of COVID-19 would be shorter than that of SARS. For both SARS and COVID-19,  
215 their infected curves have the lowest semi-saturation period (Fig 4 EJ), while cured  
216 curves have the highest which is in keeping with the common sense that the cure is  
217 always later than infection and death. For both SARS in Beijing and COVID-19 in  
218 Hubei, the  $k$  values of death case curves were lowest compared with curves of  
219 infected cases and cured cases. For both SARS and COVID-19, the  $k$  value of infected  
220 case curve was larger than that of cured curve.

221

222 ----- Insert Figure 4 near here -----

223

## 224 Discussion

225 We modeled the epidemiologic characteristics of 2019-nCoV for 31 provinces in  
226 mainland China based on the data from Jan 20 to Feb 19 during the outbreak.  
227 Furthermore, we compared these characteristics with SARS in 2003 using the same  
228 model framework.

229

230 Compared with dynamical models such as SIR (Susceptible – Infectious - Recovered)  
231 or SEIR (Susceptible – Exposed - Infectious - Recovered), our model is relatively  
232 simple but robust. We used a differential equation in logistic form to describe the  
233 dynamics of the infected, dead and healed cases caused by COVID-19 respectively  
234 (See Methods). The increasing of the population no matter in which type of case  
235 (infection, death or recovery) has multiple reasons. When the outbreak of COVID-19  
236 originated in Wuhan, it did not cause too much concern at the beginning so that it  
237 leads to many infections and starts to export potential exposed cases to other  
238 provinces. But as the successive closure of the cities in Hubei province, and isolation  
239 of the community in all the provinces in mainland China, the virus spread tendency  
240 was weakened. In our model, we also considered these measures taken by the  
241 government. We rewrote Eq. (1) to Eq. (9).

$$242 \quad \frac{dI(t)}{dt} = \alpha I(t) - rI^2(t) \quad (9)$$

243 where  $\alpha$  is growth rate of the number of infections;  $I(t)$  is cumulative number of  
244 infections at time  $t$ ;  $r$  could be understood as a parameter to reflect the effect of  
245 measures aiming at controlling the virus spread. Here we have decomposed  $I(t)$  into  
246 two parts. The first part  $\alpha I(t)$  could represent the rapid infections growth in the early  
247 phase of the outbreak without any measures for epidemic control. The second part  
248  $rI^2(t)$  illustrates the slower infections growth after taking effective isolation  
249 measures for epidemic control. For death cases, this idea still holds. But things are  
250 different for cured case. In contrast, at the early phase of the outbreak, the number of  
251 cured cases increase slowly. But once there are effective treatments, the number of

252 cured cases will rise up quickly. The growth of cured cases would again be slow since  
253 the number of patients is limited. All these facts were considered in our highly  
254 summarized model (Eq. (8)) in a macroscopic scale. All data sets have a R-square  
255 greater than 0.95 when fitted by our framework, so that we could further discuss some  
256 intrinsic mechanisms for the growth rules of infection, death and recovery through the  
257 parameters estimated from the model (See Methods).

258

259 The value of  $k$  from Eq. (8) is a critical parameter.  $k$  represents how rapid the number  
260 of infected cases get to the saturation. We found the variability of  $k$  values among  
261 fitted curves of infected COVID-19 cases in 30 provinces (Fig 2 AB) is large  
262 (Ranging from 0.19 to 0.38). It is clear to show the real situation in each province and  
263 this guides us to be more careful against the epidemic. Further we compared this  
264 characteristics between COVID-19 and SARS in 2003 in typical provinces (Fig 3 AB).  
265 We noted that  $k$  values of fitted COVID-19 infection curves in 30 provinces of  
266 mainland China (exclude Tibet) are larger than that of SARS in Beijing 2003 (Fig 3  
267 C), which suggests the time cost of this epidemic will be shorter compared with  
268 SARS in 2003. This result may related to several aspects. Firstly, it may due to the  
269 basic spread property of the virus. Besides, the isolation policy of all the provinces in  
270 mainland China was more strictly enforced than it was in 2003. Moreover, the  
271 improved clinical skills today may also helpful for shortening the duration of  
272 epidemic.

273

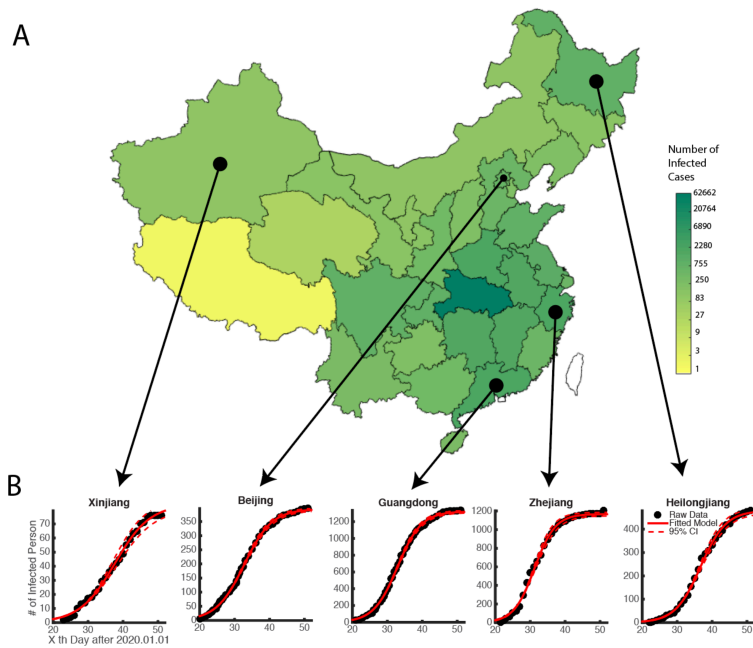
274 Furthermore, we analyzed semi-saturation period (Fig 4 EK) ( $t_0$  in Eq. (8)) and make  
275 comparison in three types of cases (infection, death and recovery) (Fig 4 A-C G-I).  
276 Our model could fit them very well ( $R\text{-square} \geq (R - s)$ ).  $k$  values for the fitted  
277 curves of the three types of cases caused by COVID-19 were larger than that of SARS,  
278 while  $t_0$  values were the opposite. This again suggests that the epidemic of COVID-19  
279 will be quicker to pass away than SARS. For both SARS and COVID-19, fitted

280 curves of infections have lowest semi-saturation period (Fig 4 EK), while fitted  
281 curves of cured cases have the highest semi-saturation period, which follows the  
282 common sense that the cure was always later than infection and death since it costs  
283 time to get on the research of effect plans to cure the patients. For both SARS and  
284 COVID-19, the  $k$  value of infected case curve was larger than that of cured curve,  
285 while that of death case is the lowest.  
286

287 **References**

- 288 Drosten C, Günther S, Preiser W, et al. Identification of a novel coronavirus in  
289 patients with severe acute respiratory syndrome. *N Engl J Med* 2003; 348: 1967–76.
- 290 Holshue ML et al. (2020) First Case of 2019 Novel Coronavirus in the United States.  
291 *N Engl J Med*.
- 292 Huang C et al. (2020) Clinical features of patients infected with 2019 novel  
293 coronavirus in Wuhan, China. *Lancet*.
- 294 HUANG De-sheng, GUAN Peng, ZHOU Bao-sen. Fitness of morbidity and  
295 discussion of epidemic characteristics of SARS based on logistic models[J]. *Chinese*  
296 *Journal of Public Health*, 2003, 19(6): 1-2.
- 297 Hui DS, E IA, Madani TA, Ntoumi F, Kock R, Dar O, Ippolito G, McHugh TD,  
298 Memish ZA, Drosten C, Zumla A, Petersen E (2020) The continuing 2019-nCoV  
299 epidemic threat of novel coronaviruses to global health - The latest 2019 novel  
300 coronavirus outbreak in Wuhan, China. *Int J Infect Dis* 91:264-266.
- 301 Ki M, nCo VT (2020) Epidemiologic characteristics of early cases with 2019 novel  
302 coronavirus (2019-nCoV) disease in Republic of Korea. *Epidemiol Health*:e2020007.
- 303 Ksiazek TG, Erdman D, Goldsmith CS, et al. A novel coronavirus associated with  
304 severe acute respiratory syndrome. *N Engl J Med* 2003; 348: 1953–66.
- 305 Kuiken T, Fouchier RAM, Schutten M, et al. Newly discovered coronavirus as the  
306 primary cause of severe acute respiratory syndrome. *Lancet* 2003; 362: 263–70.
- 307 Li Q et al. (2020) Early Transmission Dynamics in Wuhan, China, of Novel  
308 Coronavirus-Infected Pneumonia. *N Engl J Med*.
- 309 Shigemura J, Ursano RJ, Morganstein JC, Kurosawa M, Benedek DM (2020) Public  
310 responses to the novel 2019 coronavirus (2019-nCoV) in Japan: mental health  
311 consequences and target populations. *Psychiatry Clin Neurosci*.
- 312 Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y,  
313 Wang Y, Liu X, The compound Logistic model used to describe epidemic situation  
314 dynamics of SARS in Beijing, *JOURNAL OF CHINA JILIANG UNIVERSITY*,

315 2005  
316 WHO. Summary of probable SARS cases with onset of illness from 1 November  
317 2002 to 31 July 2003. Dec 31, 2003.  
318 [https://www.who.int/csr/sars/country/table2004\\_04\\_21/en/](https://www.who.int/csr/sars/country/table2004_04_21/en/) (accessed Jan 19, 2020).  
319 Zhao Y, Li Y, Wang X, Peng Z (2020) Clinical Characteristics of 138 Hospitalized  
320 Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. JAMA.  
321 Zhou P et al. (2020) A pneumonia outbreak associated with a new coronavirus of  
322 probable bat origin. Nature.  
323 Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R,  
324 Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W, China Novel  
325 Coronavirus I, Research T (2020) A Novel Coronavirus from Patients with Pneumonia  
326 in China, 2019. N Engl J Med.  
327



328

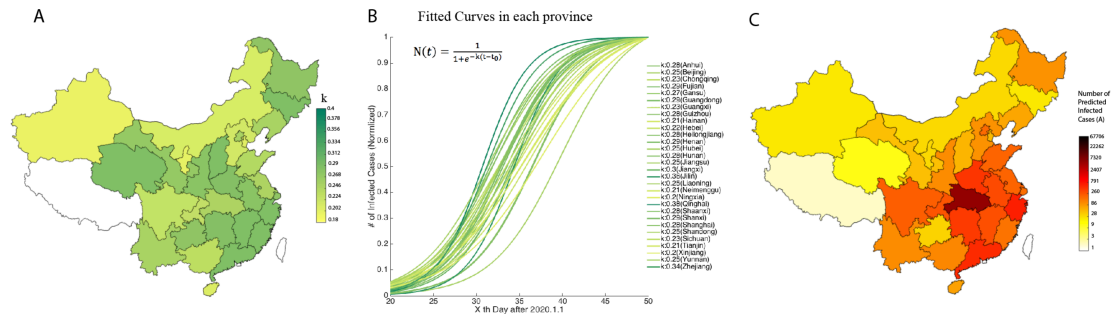
329 **Figure 1 Spatial distribution of the number of COVID-19 infected**  
330 **cases in mainland China.**

331 Panel A is the spatial distribution of the number of COVID-19 infected cases of  
332 provinces in mainland China.

333 Panel B takes five provinces (Xinjiang, Beijing, Guangdong, Zhejiang and  
334 Heilongjiang) as example to show the time series of infected cases and corresponding  
335 fitted curve. The horizontal axis is the  $x_{th}$  day after Jan 1, 2020. And the vertical axis  
336 denotes the number of infected cases in corresponding province. The black dots in  
337 each panel are raw data of the cumulative infected cases. The black dots in  
338 each panel are raw data of the cumulative infected cases. The red line is the fitted  
339 curve by our descriptive model. The dashed red lines show the 95% confidence  
340 interval of the fitted curves.

340

341



342

343 **Figure 2 Spatial distribution of increasing rate ( $k$ ) and the**  
344 **normalized fitted curves of each province of mainland China**  
345 **(exclude Tibet)**

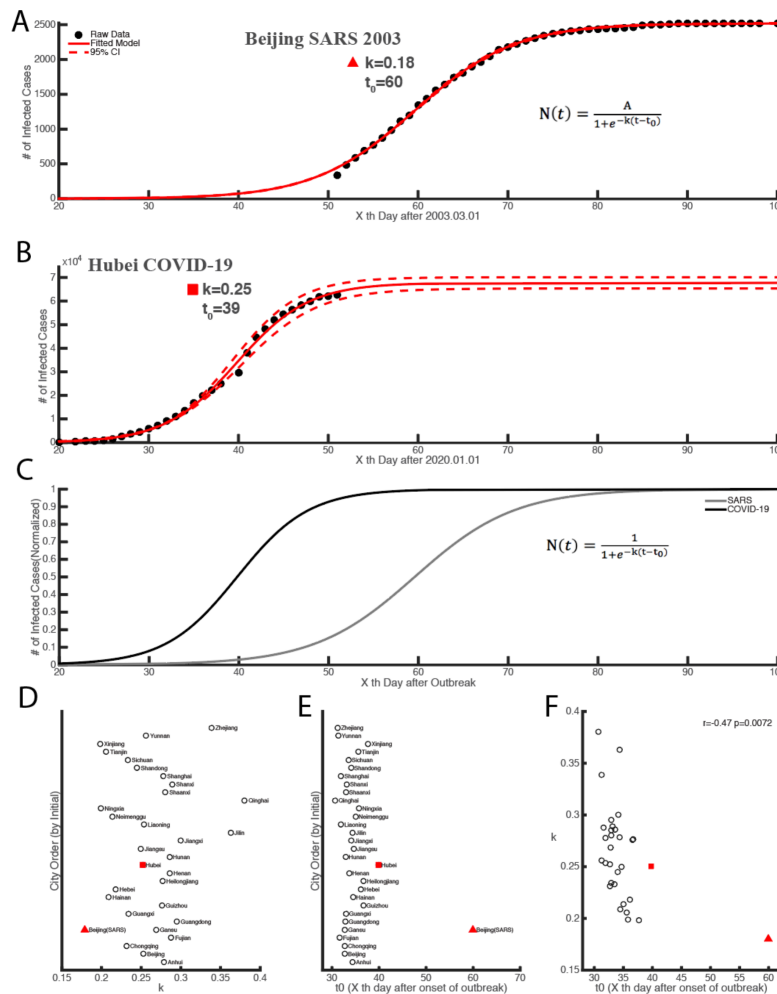
346 Panel A demonstrates the spatial distribution of parameter  $k$  of provinces in mainland  
347 China.

348 Panel B shows the fitted curves of infected cases in 30 provinces. The vertical axis is  
349 the number of infected cases (normalized by dividing the maximum number of  
350 infected cases in each province). The color of each province is consistent with Panel  
351 A.

352 Panel C is the spatial distribution of the predicted maximum number (parameter A in  
353 our model) of COVID-19 infected cases of provinces

354





355

356 **Figure 3 Comparison of intrinsic epidemic rules between COVID-19**  
 357 **in Wuhan and SARS in Beijing (2003)**

358 Panel A shows the raw data and fitted curve of SARS cases in Beijing. The horizontal  
 359 axis is the  $x_{th}$  day after March 1, 2003. The vertical axis denotes the cumulative  
 360 number of SARS cases in Beijing. The black dots are the raw data of the SARS cases.  
 361 The red line is the fitted curve by our descriptive model. The dashed red lines are the  
 362 95% confidence interval of the fitted curves.

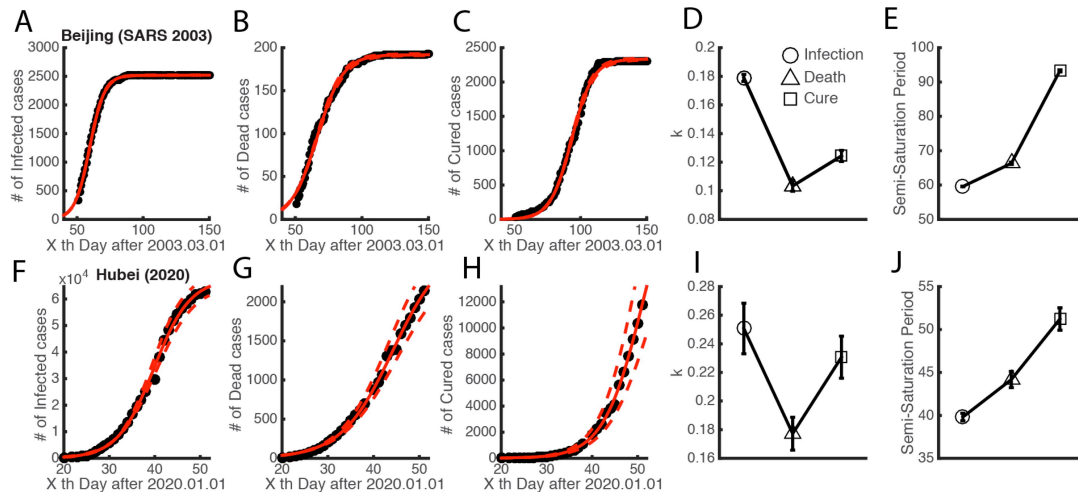
363 Panel B shows the raw data and fitted curve of COVID-19 cases in Hubei province.  
 364 The horizontal axis is the  $x_{th}$  day after Jan 1, 2020. The vertical axis denotes the  
 365 cumulative number of COVID-19 cases in Hubei. The black dots are raw data of the  
 366 COVID-19 cases. The red line is the fitted curve by our descriptive model. The  
 367 dashed red lines are the 95% confidence interval of the fitted curves.

368 Pannel C shows the fitted curves of infected SARS cases (gray curve) and COVID-19  
369 (black curve), which are normalized by dividing the maximum number of infected  
370 cases in each province.

371 Panel D illustrates the  $k$  value (increasing rate of infected cases) of fitted curves of  
372 infected COVID-19 cases in each province in mainland China, where Hubei is  
373 represented by the solid red square (the same in D and E).  $k$  value of fitted curves of  
374 infected SARS cases in Beijing is represented by the solid red triangle (the same in D  
375 and E).

376 Panel E illustrates the  $t_0$  (semi-saturation period) of fitted curves of infected  
377 COVID-19 cases in each province in mainland China, as well as  $t_0$  of fitted curves of  
378 infected SARS cases in Beijing. The scatter plot of  $k$  and  $t_0$  is shown in Panel F,  
379 indicating significant negative correlation between  $k$  and  $t_0$ .

380



381

382 **Figure 4 Intrinsic rules of infection, death and recovery of COVID-19**  
 383 **and SARS2003**

384 Panel A-C show the raw data and fitted curve of SARS cases in Beijing. Their  
 385 horizontal axes are the  $x_{th}$  day after Mar 1, 2003. The black dots in Panel A, B and C  
 386 are raw data of the SARS infected cases, dead cases and cured cases, respectively.  
 387 The red lines are the fitted curves by our descriptive model. The dashed red lines are  
 388 the 95% confidence interval of the fitted curves.

389 Panel D and Panel E respectively shows the  $k$  value and the  $t_0$  of fitted curves of three  
 390 types cases caused by SARS (the circle, triangle and square denotes infected, dead  
 391 and cured cases, respectively, and the same in Panel I and J). The error bars show the  
 392 95% confidence interval of the parameters.

393 Panel F-H show the raw data and fitted curve of COVID-19 cases in Hubei. Their  
 394 horizontal axes are the  $x_{th}$  day after Jan 1, 2020. The black dots in Panel F, G and H  
 395 are raw data of the COVID-19 infected cases, dead cases and cured cases,  
 396 respectively. The red lines are the fitted curves by our descriptive model. The dashed  
 397 red lines are the 95% confidence interval of the fitted curves.

398 Panel I and Panel J respectively shows the  $k$  value and the  $t_0$  of fitted curves of three  
 399 types cases caused by COVID-19, which is same with D and E.