Title: **The Enlightenment of mainland China's epidemic situation to the world through the intrinsic growth rule of infected and cured cases with COVID-19**

Chuanliang Han[1△], Yimeng Liu[2,3,4△], Jiting Tang[2,3,4] , Yuyao Zhu[2,3,4] and Saini Yang[2,3,4*]

[1]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

[2]Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing 100875, China

[3]State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

[4]Academy of Disaster Reduction and Emergency Management Ministry of Emergency Management and Ministry of Education, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

[△]C.L.H and Y.M.L contributed equally to this work

***Correspondence authors:**

Saini Yang,

Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education & State Key Laboratory of Earth Surface Processes and Resource Ecology & Academy of Disaster Reduction and Emergency Management Ministry of Emergency Management and Ministry of Education, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

Email: yangsaini@bnu.edu.cn

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

Acknowledgements: This study is sponsored by the National Key Research and Development Program of China (2018YFC1508903), the National Natural Science Foundation of China (41621061) and the support of International Center for Collaborative Research on Disaster Risk Reduction (ICCR-DRR).

Abstract

The novel coronavirus disease (COVID-19) emerged at the end of 2019 has been controlled in mainland China so far, while it is still spreading all over the world and causing global panic. When the pandemic will end is an issue of great concern. A logistic model depicting the growth rules of infected and cured cases in mainland China may shed some light on this question. We extend this model to 31 countries outside China experiencing serious COVID-2019 outbreaks. The model well explained data from each country in our study (R-Square ≥ 0.95). For the infected case, the semi-saturation period of these countries are ranging from 63 to 170 (Mar 3 to Jun 18). For all countries in our study, the increasing rate (k) and semi-saturation period (t_0) of the fitted curves are negatively correlated for both infected and cured cases. The increasing rate of infected case is positively correlated with that of cured case, so is the semi-saturation period. We used a simple linear model to bridge the relationship between the increasing rules of infected and cured case. According to the connection of infected and cured growth rules which is identified from cases in China, we predicted that the semi-saturation period of the cured case outside China ranges from 82 to 196 (Mar 22 to July 8) for cured cases. More importantly, we found there exists strong positive correlation between the semi-saturation period (t_0) of infected cases and the timing of government's response, which could provide strong evidence for the effectiveness of rapid epidemic prevention in various countries.

Introduction

In December 2019, a series of cases with unknown cause of pneumonia appeared in Wuhan, the capital of Hubei Province in mainland China (WHO[a]). Deep sequencing analysis later from lower respiratory samples confirmed a novel coronavirus, which was firstly named as 2019 novel coronavirus (2019-nCoV) on Jan 12, 2020 and may originate from bat (Zhou et al., 2020). The coronavirus (SARS-CoV-2) has been officially renamed as COVID-19 on February 12, 2020. Human to human transmission through COVID-19 has been confirmed not only in China (Wang et al., 2020; Li et al., 2020; Zhu et al., 2020; Hui et al., 2020), but in dozens of countries around world such as Republic of Korea (Choi et al., 2020; Ki et al., 2020; Shim et al., 2020), Italy (Livingston et al., 2020; Spina et al., 2020; Rosenbaum, 2020) and Iran (Tuite et al., 2020; Abdi, 2020; Zhuang et al., 2020). So far, the epidemic situation in mainland China has been effectively controlled since there are only sparse daily local new cases since Mar 19, 2020, and the cure rate in mainland China has risen up to 93.51% by April 13 (Chinese Center for Disease Control and Prevention, CCDC). However, the epidemic situation outside China is much more serious than before. As of Apr 13, there have been a total of 1788665 confirmed cases of COVID-19 globally, and the death and cured rate are 6.35% and 20.45 % respectively (Sina.com). When the inflection point will come for both infection and cure outside China remains unclear. The hope and lessons for COVID-19 control from China is necessary and worth to quantify (Azman & Luquero, 2020).

We applied a descriptive model that has been proved robust and stable (Han et al., 2020) to analyze the data on the COVID-19 global cases (infected and cure cases in mainland China, and infected cases countries outside China). We then obtained the parameter relationship between parameters of infected and cure cases in mainland China. Further, we used that relationship to map parameter space of infected case to the parameter space of cured cases for 31 countries globally. In the end, we tried to explore the relationship

between the model parameters and governmental control measures.

Method

Sources of Data:

The cumulative number of confirmed and cured COVID-19 cases in mainland China was obtained from the National Health Commission of China, and the provincial Health Commission of 30 provincial administrative regions (excluded Tibet, because the only one confirmed infected COVID-19 case in Tibet has been cured on Feb. 12) in mainland China (January 10 to March 19, 2020), which is publically available. All cases were laboratory confirmed following the published standards made by the National Health Commission of China ($CCDC_a$). The basic test procedure has been described in detail in previous work (Zhou et al., 2020; Huang et al., 2020). This dataset was partially analyzed in an initial work on this topic (Han et al., 2020).

We selected 31 countries outside China with serious coronavirus epidemic situation in our analyses (January 10 to Apr 13, 2020), whose population counts for 39.4% of the world's population (57.8% if includes China) and 81.6% of the world's infected COVID-19 cases (86.2% if includes China) (WHO_a). The data of COVID-19 cases in countries outside China were obtained from the situation reports on the official website of World Health Organization (WHO_a), which is publicly available. The data used in this study include the cumulative number of reported laboratory-confirmed COVID-19 cases. Countries included in our study are Republic of Korea, Japan, Australia, Singapore, which are in Western Pacific Region; Italy, Spain, France, Germany, Switzerland, the United Kingdom, Netherlands, Sweden, Denmark, Austria, Belgium, Portugal, Czechia, Finland, Ukraine, Slovakia, Bulgaria, Lithuania, which are in European Region; India in South-East Asia Region, Islamic Republic of Iran, Lebanon, which are in Eastern Mediterranean Region; United States of America, Canada, Brazil, Mexico, which are in

America; South Africa, Ethiopia, which are in Africa. All the laboratory-confirmed cases are determined according to the WHO standard.

Governments of these countries all declared a state of national emergency, wartime, or a blockade of their borders as a result of the COVID-19 outbreak. In this study, we collected the time point of government declaration from mainstream authoritative media of each country we selected.

Epidemic curve modeling

In this work, we used a sigmoid-formed descriptive model (Equation (1)), which is a derivation from logistic form differential equations (Han et al., 2020).

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

where $N(t)$ is the general form of the cumulative numbers of infected or cured patients at time t. A denotes the maximum number of infections or cures, k represents incidence growth rate, t_0 is the inflection point of the sigmoid curve. This descriptive model has been verified by not only the infected, death and cured cases of COVID-19, but also the data of SARS in 2003 (Han et al., 2020). For infected case, there are three parameters (A, k, t_0) in the model, and for cured case, we fixed A to the subtraction of the maximum number of cumulative infections and death, which obeys the basic biological fact. It is worth mentioning that in our model, t_0 is the inflection point in mathematical definition. In this paper, we assumed that t_0 is the time of inflection point of epidemic situation in a region.

We processed the data and modeling with custom scripts on MATLAB (the Math Works). We adopted the nonlinear least square (NLS) algorithm for data fitting and parameter estimation.

136 **Results**

137 With the cumulative number of confirmed and cured cases in 30 provinces of mainland
138 China as well as in 31 countries and above-mentioned model, all data could be well
139 explained (R^2>0.95) by the descriptive model (See Methods, Fig. 1A-C, Fig. 4A-C).
140

141 **Relationship between infected and cured cases in mainland China**

142 The epidemic situation of COVID-19 has been controlled in China ($CCDC_b$), with a cure
143 rate of 93.51% ($CCDC_a$). Time series of infected and cured cases and their fitted curves
144 of three example provinces (provinces of Anhui, Heilongjiang, Hebei) are shown in Fig.
145 1A-C. By observing the fitted curves of all provinces in mainland China (Supplementary
146 Figure S1), we found the increasing rate (k) of infected case is larger than that of cured
147 case and the semi-saturation period (t_0) of infected cases is shorter than that of cured
148 cases.

149 As shown in Figure 2, the parameter space of k and t_0 are separated for infected (red
150 circle) and cured cases (blue circle), and there is a significant negative correlation (for
151 infected case, r=-0.46, p=0.0103; for cured case, r=-0.56, p=0.0012; Pearson correlation).
152 As we compared these two parameters of infected and cured cases respectively, we
153 found infection case has larger (t=6.9136, $p<10^{-7}$, right-tailed test) increasing rate (Fig.
154 2B) but smaller (t=-37.271, $p<10^{-25}$, left-tailed test) semi-saturation period (Fig. 2C) than
155 that of cured cases. The increasing rate of infected and cured case is significant
156 positively correlated (r=0.38, p=0.0359, Pearson Correlation), so is the semi-saturation
157 period of infected and cured case (r=0.66, p=0.0001, Pearson Correlation). Each of them
158 could be fitted to a linear model, with a slope and intercept of 0.36 (CI: [0.0268, 0.698])
159 and 0.1 (CI: [0.014, 0.191]), respectively (panel B), and that of panel C is 1.06 (CI: [0.59,
160 1.54]) and 15.36 (CI: [-0.85, 31.56]).
161

162 **Intrinsic growth rules of COVID-19 infections outside China**

163 With the dataset of infected case in mainland China, there shows time stability of this

model (Fig. 3). An example of Fujian province is shown in Figure 3A. It is clear that as the increase of length of time-series data, the fitting curves converges better to the actual number of cases (black dots).

We defined an index, the minimum time length (MTL), to measure how much data is necessary to produce stable output. This index is the time length with which the change of parameter value (k or t_0) is less than 5% in two consecutive days. With the data of 30 provinces in China, we found that the MTL for increasing rate (k) is 15 days (Fig. 3B) and for semi-saturation period (t_0), is 26 days (Fig. 3C). This result provides some hint on how much data is necessary to measure the epidemic situation.

After identifying the relationship between the increasing rate and the semi-saturation period, we applied this relationship to the COVID-19 cases from 31 countries. We applied the descriptive model to fit the data of the cumulative infected cases in each of the 31 countries. Figure 4 illustrates three countries' time-series data of infection (Canada, Germany and Iran, Fig. 4A-C) (See Supplementary Figure S2 for all 31 countries' plots). All data series of infected case could be well explained by our model (R^2>0.95). The average change rate of these parameters is already lower than 5% (Fig 4 D-F) which indicates the parameters we estimated from the model is stable. We noticed that the increasing rate of 31-country cases does not have a significant difference (t=1.197, p=0.237 two-tailed test) with that of Chinese provinces, but the 31-country cases have significantly longer (t=-17.4819, $p<10^{-20}$, right-tailed test) semi-saturation period. The numbers indicated that the t_0 of infections in 31 countries is averaged at 98 (std: 20.5), and the date matched to which is around Apr 7 (ranging from Mar 3 to Jun 18).

Prediction for growth rules of cured COVID-19 cases and its relation to the timing of governmental control measures

By observing the increasing rate (k) and semi-saturation period (t_0) of infected case in 31

countries, we found there was a weak negative correlation (r=-0.63, p=0.0001, Pearson Correlation) (Fig. 5) between these two parameters, which shows similar tendency with results of provinces in mainland China (Fig. 2B).To predict the increasing rate (k) and inflection point (t_0) of the cured COVID-19 cases outside China, we mapped the parameter space of infected cases into the parameter space of cured cases, based on the parameter relationship between infected cases and cured cases obtained from China (Fig. 2BC). For the 31 countries, the increasing rate ranges from 0.08 to 0.45. The mean t_0 of the fitted cured cases curves is 119.71 (std: 21.73), which means the mean semi-saturation period (t_0) for these 31 countries will come around Apr. 29, 2020 (ranging from Mar 22 to July 8).

To explore the practical implication of these parameters, we took the timing of governmental emergency control measures related to COVID-19 as a variable to analyze its relationship with the key parameters (Fig. 6). We found the Semi-Saturation Period (SSP) showed a significant positive correlation (r=0.73, p<0.0001, Spearman correlation) with the timing of governmental control measures (Fig. 6 BE), and k showed a marginal significant negative correlation (r=-0.3, p=0.0925, Spearman correlation), while the increasing rate or the maximum of infected number do not show strong correlation (r=0.02, p=0.91, Spearman correlation) with the control measures' timing (Fig. 6 ACDF). China released the earliest national control measures to prevent the spread of COVID-19 (on Jan. 23) and its t_0 is the lowest (Figure 5, red circle). We also noticed that the timing of governmental emergency control measure is significantly earlier than SSP (Fig. 6B) (t=2.99, p=0.0043, two-tailed test) in time, which indicates that most of the selected countries have taken measures at the early phase of the outbreak.

Discussion

We compared the relationship between growth characteristics of infected and cured case of 30 provinces in mainland China based on the data from Jan 20 to Mar 29, 2020, which are positively correlated (Fig. 2 BC). Further, we explained the data from 31 countries using the same model (See Methods). We predicted their increasing rate (k) and semi-saturation period (t_0) of cure case based on the knowledge of data from China. It is worth mentioning that there is a strong correlation between these parameters and the timing of governmental policy on COVID-19.

Comparison with previous work

Different from models like SIR or SEIR applied in COVID-19 (Li et al., 2020; Wang et al., 2020; Tian et al., 2020), our model is relatively simple, but robust. This work is an extension of our previous work (Han et al., 2020), we have verified that our model well explains the data in various geographic space of mainland China, as well as provided robust results with different lengths of time series (between 30 and 40 days).

This model is able to capture the features of macroscopic dynamic of the epidemics worldwide (R^2>0.95) (Fig. 4) and we quantitatively investigated the minimum data length to produce stable output (Fig. 3). By now, the domestic infected cases in China is very limited, which is consistent with our prediction results. The cumulative number of infected and cured cases are almost saturated, and this means the spread of the virus has been effectively controlled, thus we believe that the parameters estimated in this work for Chinese provinces are reliable (Fig. 1A-C). The characteristics of two curves can depict the dynamic macroscopic features of the epidemic situation. Moreover, we extended this model to the data in each of these 31 countries (Fig 4 A-C), and found the sigmoid model could also fit the data very well (R^2>0.95) (Fig 4 A-C).

Relationship among model parameters and the prediction of the cured cases

We noticed that semi-saturation period of infected and cured cases in China showed a

positive correlation (p-value = 0.0001). This phenomenon obeys the common sense, since the cure is always later than infection when a new virus comes. Although the positive correlation between increasing rate of infected and cured cases is significant (p-value = 0.0359), we noticed the correlation is much lower than that of semi-saturation period. This result is reasonable, since the reasons cause the ascending increasing rate of infected case are totally different from that of cured case. So that correlation may not be that strong. As we compared the k and t_0 in any specific case no matter in mainland China or worldwide region (Fig. 5), we found significant negative correlations. This is consistent with our initial work (Han et al., 2020). For negative correlation of k and t_0 in cured cases, it could also be explained as when the increasing rate (k) of the cured case is larger, it will be faster to cure most patients.

Suggestion for governments around the world on the prevention and control of COVID-19

More importantly, we found there is a strong correlation between the timing of governmental control measures for COVID-19 in 31 countries and the semi-saturation period estimated (Fig 6 BE). This indicates the early implementation of the government's prevention and control policy effectively shortened the turning point of the epidemic. Taken China as an example, when the COVID-19 outbreak in Wuhan, the strict isolation policy was announcement on Jan. 23, and the inflection point there was around Feb 9, 2020, the time lag was 17 days. Another result showed that the time of governmental policy is significantly shorter than the inflection time point of infection, which is also a strong reference for governments of all countries to take control measures as early as possible.

Reference

Sina, https://news.sina.cn/zt_d/yiqing0121

WHO_a, https://www.who.int/docs/default-source/coronaviruse/situation-reports

$CCDC_a$ Prevention, C. C. f. D. C. a. Prevention and control measures of COVID-19, <http://www.nhc.gov.cn/jkj/s3577/202002/a5d6f7b8c48c451c87dba14889b30147/files/3514cb996ae24e2faf65953b4ecd0df4.pdf> (2020).

$CCDC_b$ http://www.nhc.gov.cn/xcs/yqfkdt/gzbd_index.shtml

Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J (2020) Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science.

Zhou P et al. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature.

Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y, Zhao Y, Li Y, Wang X, Peng Z (2020) Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. JAMA.

Li Q et al. (2020) Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med.

Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W, China Novel Coronavirus I, Research T (2020) A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med.

Hui DS, E IA, Madani TA, Ntoumi F, Kock R, Dar O, Ippolito G, McHugh TD, Memish ZA, Drosten C, Zumla A, Petersen E (2020) The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 91:264-266.

Choi SC, Ki M (2020) Estimating the reproductive number and the outbreak size of Novel Coronavirus disease (COVID-19) using mathematical model in Republic of Korea. Epidemiol Health:e2020011.

Ki M, nCo VT (2020) Epidemiologic characteristics of early cases with 2019 novel coronavirus (2019-nCoV) disease in Republic of Korea. Epidemiol Health:e2020007.

Shim E, Tariq A, Choi W, Lee Y, Chowell G (2020) Transmission potential and severity of COVID-19 in South Korea. Int J Infect Dis.

Abdi M (2020) Coronavirus disease 2019 (COVID-19) outbreak in Iran; actions and problems. Infect Control Hosp Epidemiol:1-5.

Tuite AR, Bogoch, II, Sherbo R, Watts A, Fisman D, Khan K (2020) Estimation of Coronavirus Disease 2019 (COVID-19) Burden and Potential for International Dissemination of Infection From Iran. Ann Intern Med.

Zhuang Z, Zhao S, Lin Q, Cao P, Lou Y, Yang L, He D (2020) Preliminary estimation of the novel coronavirus disease (COVID-19) cases in Iran: A modelling analysis based on

overseas cases and air travel data. Int J Infect Dis.
Livingston E, Bucher K (2020) Coronavirus Disease 2019 (COVID-19) in Italy. JAMA.
Rosenbaum L (2020) Facing Covid-19 in Italy - Ethics, Logistics, and Therapeutics on the Epidemic's Front Line. N Engl J Med.
Spina S, Marrazzo F, Migliari M, Stucchi R, Sforza A, Fumagalli R (2020) The response of Milan's Emergency Medical System to the COVID-19 outbreak in Italy. Lancet 395:e49-e50.
Han C, Liu Y, Yang S (2020) Intrinsic growth rules of patients infected, dead and cured with 2019 novel coronavirus in mainland China. medRxiv:2020.2002.2023.20024802.
Zhou P et al. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature.
Huang C et al. (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet.
Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J (2020) Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science.
Wang H, Wang Z, Dong Y, Chang R, Xu C, Yu X, Zhang S, Tsamlag L, Shang M, Huang J, Wang Y, Xu G, Shen T, Zhang X, Cai Y (2020) Phase-adjusted estimation of the number of Coronavirus Disease 2019 cases in Wuhan, China. Cell Discov 6:10.
H. Tian et al., Science 10.1126/science.abb6105 (2020).
Azman AS, Luquero FJ (2020) From China: hope and lessons for COVID-19 control. Lancet Infect Dis.

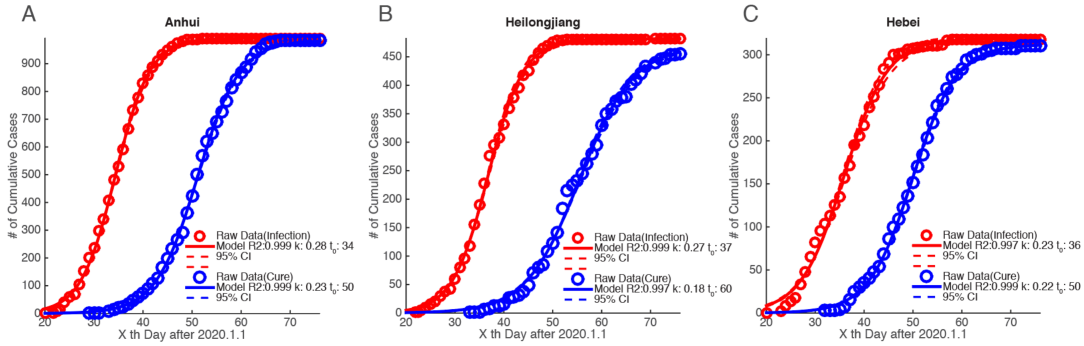


Figure 1 Example provinces for the time series of infected and cured cases with corresponding fitted curve in mainland China

A-C shows three example cases in three provinces (Anhui, Heilongjiang and Hebei). The horizontal axis in each panel denotes the xth day after January 1, 2020. The vertical axis in each panel is the cumulative number of infected (red circles) and cured case (blue circles). The red and blue line are the fitting curves by the sigmoid model, and the dashed red and blue curves mark the 95% confidence interval of the fitting curves.

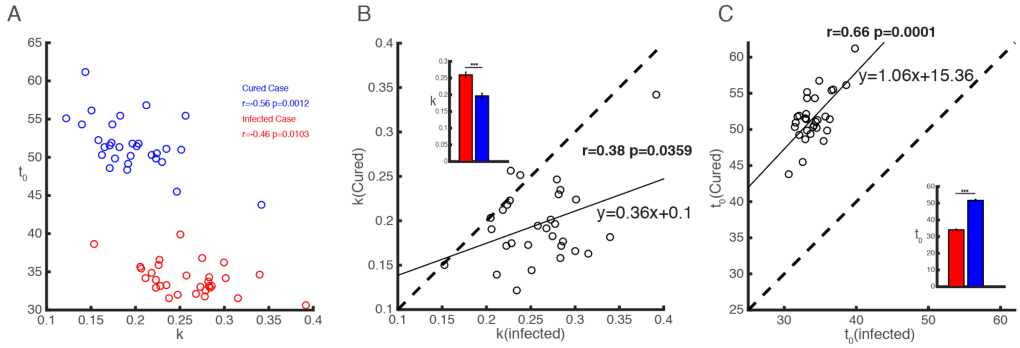


Figure 2 Relationship between intrinsic rules for infected and cured case in mainland China

Panel A shows the scatter plot of increasing rate (k) and semi-saturation period (t_0) for infected (red circles) and cured case (blue circles).

Panel B is the scatter plot of increasing rate (k) of infected and cured case (black circles), the dashed line is the diagonal line of the coordinate system (the same for Panel C). The relation between k of infected and cured cases is modeled with a linear function, which is shown as the black curve. The bar graph shown in northwest corner demonstrates the difference between *k* of infected and cured cases.

Panel C shows the scatter plot of semi-saturation period (t_0). The black curve is the linear function estimated based on the relation between infected and cured case. And the bar graph shown in southeast corner demonstrates difference between t_0 of infected and cured cases.

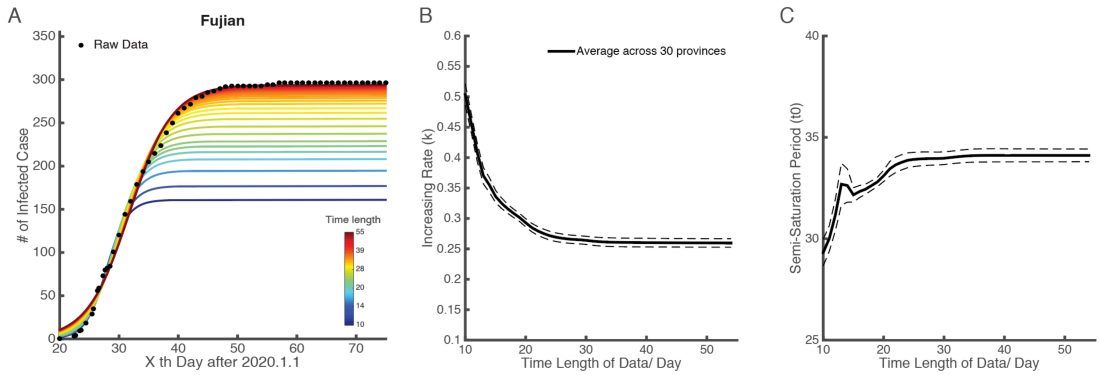


Figure 3 Time stability of logistic model in infected case of Chinese mainland

Panel A is an example case in Fujian province. The horizontal axis in each panel denotes the x th day after January 1, 2020. The vertical axis in each panel is the cumulative number of infected (black dots). The lines with different colors are the fitting curve by the logistic model using data with different time length as shown in color bar in the southeast corner. Panel B and C illustrates how estimated parameters (k and t_0) changes with increase of the data's time length. The black curve shows the average of increasing rate (k) across 30 provinces in mainland China changing with time length of data and the dashed line marks its standard error.

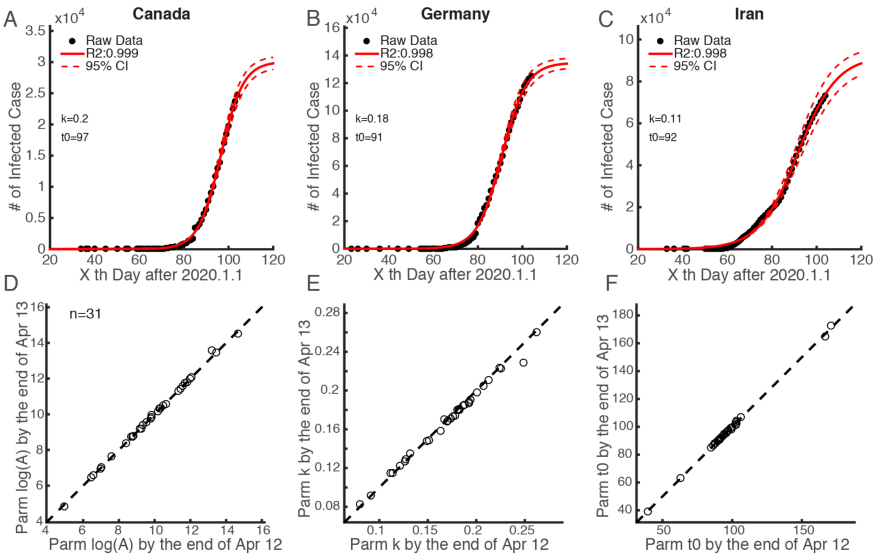


Figure 4 Example countries for the time series of infected cases with corresponding fitted curve around world and its temporal stability

Panel A-C show infected cases and their fitted curves of Australia, Italy and Iran. The horizontal axis is the xth day after Jan 1, 2020. The vertical axis denotes the number of infected cases in corresponding country. The black dots in each panel are raw data of the cumulative number of infected cases. The red line is the fitted curves by our descriptive model. The dashed red lines mark the 95% confidence interval of the fitted curves. Panel D-F show the difference of parameters estimated from relative longer (by the end of Jun 13) and shorter (by the end of Apr 12) time series data, each dot denotes for one country.

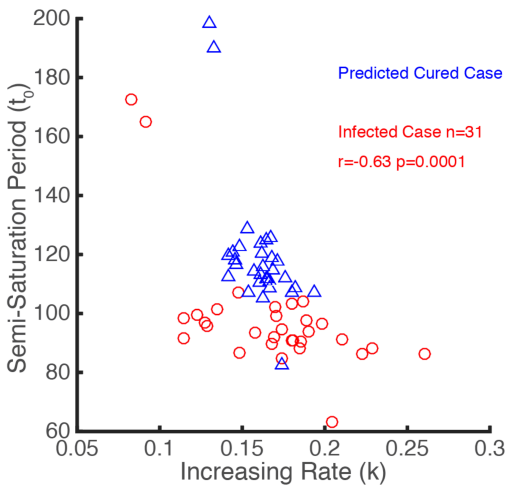


Figure 5 Prediction for intrinsic rule of cured case around world based on Chinese enlightenment

The scatter plot of increasing rate (k) and semi-saturation period (t_0) for infected cases is showed in red circles. Based on the linear model shown in Fig 2BC, a prediction of k and t_0 for cured case is plotted as blue circles.

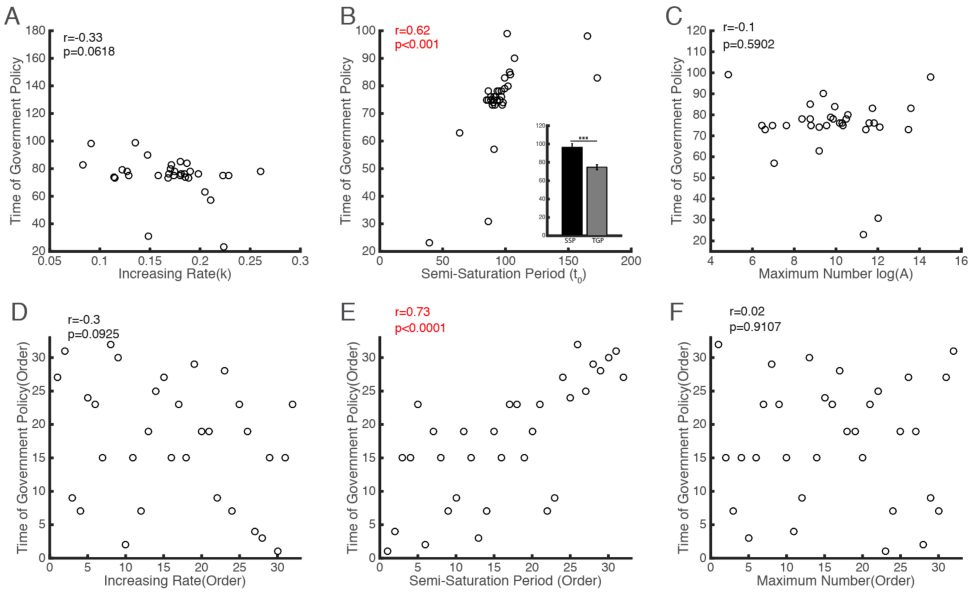
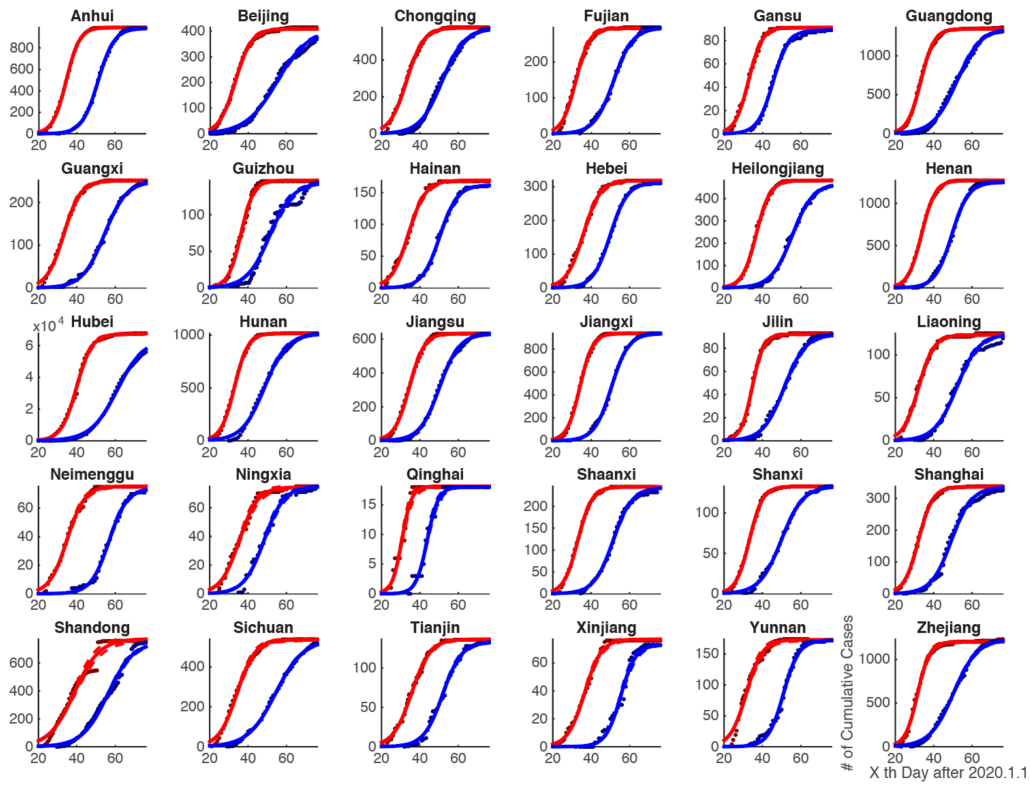


Figure 6 Relationship between time of government policy from each countries and parameters estimated from the model

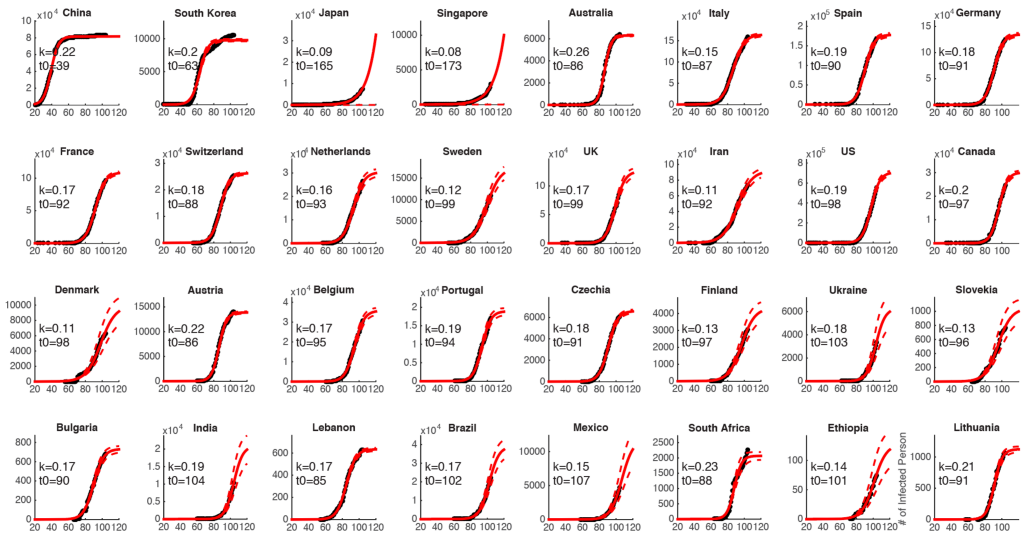
Panel A-C show the scatter plot of the time of government measures and the model parameters of infected case (k, t_0 and A, respectively). In the southeast corner of Panel B, a bar graph was shown to indicate the difference between timing of government measures (grey bar) and semi-saturation period (SAP, black bar).

D-F shows the same data of A-C but in the order form. The red circle in six panels denotes China.



Supplement Figure S1 Intrinsic growth rules of patients infected with 2019 novel coronavirus in all provinces of mainland China

The deep red and blue dots in each panel are raw data of the cumulative infected and cured cases. The red and blue line is the corresponding fitted curves by our descriptive model.



Supplement Figure S2 Intrinsic growth rules of patients infected with 2019 novel coronavirus in 31 countries around world

The black dots in each panel are raw data of the cumulative infected cases. The red line is the fitted curve by our descriptive model.