RESEARCH ARTICLE

Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis

Yaqing Fang[1] | Yiting Nie[2] | Marshare Penny[3]

[1]School of Public Health, Hangzhou Medical College, Hangzhou, China

[2]Graduate School, Harbin Medical University, Harbin, China

[3]Department of Public Health Sciences, California Baptist University, Riverside, California

Correspondence
Yaqing Fang, School of Public Health, Hangzhou Medical College, Hangzhou, 310053, Zhejiang, China.
Email: colourwind1035@126.com

Abstract

Using the parameterized susceptible-exposed-infectious-recovered model, we simulated the spread dynamics of coronavirus disease 2019 (COVID-19) outbreak and impact of different control measures, conducted the sensitivity analysis to identify the key factor, plotted the trend curve of effective reproductive number (R), and performed data fitting after the simulation. By simulation and data fitting, the model showed the peak existing confirmed cases of 59 769 arriving on 15 February 2020, with the coefficient of determination close to 1 and the fitting bias 3.02%, suggesting high precision of the data-fitting results. More rigorous government control policies were associated with a slower increase in the infected population. Isolation and protective procedures would be less effective as more cases accrue, so the optimization of the treatment plan and the development of specific drugs would be of more importance. There was an upward trend of R in the beginning, followed by a downward trend, a temporary rebound, and another continuous decline. The feature of high infectiousness for severe acute respiratory syndrome coronavirus 2(SARS-CoV-2) led to an upward trend, and government measures contributed to the temporary rebound and declines. The declines of R could be exploited as strong evidence for the effectiveness of the interventions. Evidence from the four-phase stringent measures showed that it was significant to ensure early detection, early isolation, early treatment, adequate medical supplies, patients' being admitted to designated hospitals, and comprehensive therapeutic strategy. Collaborative efforts are required to combat the novel coronavirus, focusing on both persistent strict domestic interventions and vigilance against exogenous imported cases.

KEYWORDS

basic reproductive number, COVID-19, data fitting, data simulation, effective reproductive number, effectiveness, intervention, SARS-CoV-2, SEIR, sensitivity analysis, transmission

1 | BACKGROUND

On 30 January 2020, the epidemic of coronavirus disease 2019 (COVID-19) was declared as a Public Health Emergency of International Concern, the highest level in the World Health Organization's (WHO) emergency response for infectious diseases. The number of cases was accelerating in China and subsequently all over the world; however, the publications of the COVID-19's potential transmission and effectiveness of government interventions were limited. It is urgent to provide more scientific information for a better understanding of the novel coronavirus and further containment of the outbreak.

2 | INTRODUCTION

An unknown virus was suspected to have emerged into the human population in Wuhan in late December 2019 through a report of several pneumonia cases by a local hospital.[1] Chinese health authorities took actions immediately and had detected a new virus relevant to the outbreak of disease, which was known as novel coronavirus (nCoV) infected pneumonia by 8 January 2020[2] and later be designated as COVID-19 by WHO.[3] The virus was named as the 2019-nCoV temporarily on 12 January 2020[1] and officially named as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) on 11 February 2020[3] by WHO.

Coronaviruses, getting their name from the viruses' vague resemblance to monarchical crowns when imaged using an electron microscope, are a large family of viruses that cause disease in mammals and birds. Coronaviruses can cause illnesses that range from the common cold to much more severe illnesses like SARS, Middle East respiratory syndrome, and COVID-19.[2]

The outbreak of COVID-19 originated from the four admitted patients with pneumonia who had been working in Wuhan Huanan seafood wholesale market, doing business in live poultry, aquatic products, and some wild animals. The now-closed market being a common factor in infections encouraged the belief that the infection may be linked with certain animals. The species that harbored the SARS-CoV-2 was probably bat, containing 96% identical at the whole-genome level.[4] The COVID-19 may include signs of fever, cough, shortness of breath and general breathing difficulties, organ failures or even death, posing a severe threat to the whole society. The WHO declared the coronavirus outbreak a global health emergency on 30 January 2020.[5] The global epidemic seemed to be spreading at an alarming rate, causing 79 824 accumulative laboratory-confirmed infections with 2870 deaths as of 29 February 2020 in China[6] and 6009 outside China,[7] so it deserved priority attention and intensive research. Chinese health authorities stated that it was likely to transmit from person to person even before any actual signs appeared, which made it especially difficult to prevent and control.[8]

However, Chinese government have attached extremely great importance and taken a range of prompt public health measures in response to the epidemic, starting by the strict exit screening measures in Wuhan on 23 January and followed by extensive medical support in Wuhan, individual going out under protective equipment (masks), school postponement, cancellation of mass gatherings, spontaneous isolated population, and so on.

Meanwhile, researchers all around the world have been scrambling to ascertain how the virus spreads and find out the effective ways to put this outbreak quickly under control. Compared to the R_0

of H1N1 (1.25)[9] and that of SARS (2.2-3.6),[10] the reproduction number of COVID-19 indicates awful potential transmission. The R_0 was estimated as 2.2,[8] 3.8,[11] and 2.68[12] by different research all round the world. The WHO published an estimated R_0 of 1.4 to 2.5.[13] R_0, the basic reproductive number, is defined as the expected number of additional cases that one case will generate, on average without external intervention. R_0 value is usually applied to reflect the potential and severity of infectious diseases. The larger R_0 is, the stronger the transmission power will be. If R_0 is less than 1, it means that the disease will gradually die out.[14] Despite all of the results of the nCoV, the epidemic kept in a constant state of change, especially after the rigorous measures have been taken by Chinese government, so it calls for more comprehensive research such as updated R_0 estimates, more appropriate index under the circumstances of intervention such as effective reproductive number (R), methodological improvements, and effect evaluation of control measures to reflect the dynamic and progress of it. In this study, we provided the latest distributions of COVID-19 in Wuhan, China, and the whole world. Besides, we adopted the model of susceptible-exposed-infectious-recovered (SEIR) to estimate the dynamics and the potential spread based on the current data of cases, to calculate the R_0 and R under different scenarios of the epidemic, and to draw preliminary conclusions about the effectiveness of government measures.

3 | METHODS

3.1 | Data

The data of COVID-19 in the study were mainly obtained from National Health Commission of the People's Republic of China, Chinese Center for Disease Control and Prevention, WHO, and various websites of Chinese government agencies, official media, as well as some previous studies as of 29 February 2020.

3.2 | Mathematical model

One of the mathematical models in epidemic dynamics, known as the "warehouse" model, has been widely applied for a long time since it was forwarded by Kermack and McKendrick in 1927. It includes several basic and improved models, such as SIR, SIS, SEIR, and so on, among which SEIR is a typical example that takes the incubation period into account.[15] The SEIR, an extensively used epidemic model, can reflect the flows of people between four states: susceptible (S), exposed (E), infectious (I), and recovered (R).
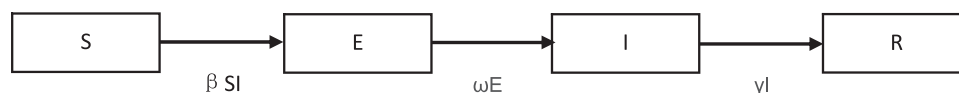
In this study, equations for the population change of each warehouse were established as follows[15]

$$
\begin{cases}
\dfrac{dS}{dt} = -\beta SI \\[2mm]
\dfrac{dE}{dt} = \beta SI - \omega E \\[2mm]
\dfrac{dI}{dt} = \omega E - \gamma I \\[2mm]
\dfrac{dR}{dt} = \gamma I
\end{cases},
$$

where $\beta = \beta_0 k$, β denoted the coefficient of infection rate, β_0 denoted the probability of infection per exposure, k denoted the frequency of exposure; $\omega = 1/T_e$, ω denoted the coefficient of migration rate of latency, T_e denoted the average latency; and $\gamma = 1/T_i$, γ denoted the coefficient of migration rate, T_i denoted the average recovery time.

The incidence in the original model was bilinear, because the model assumed that the frequency of exposure U was directly proportional to the total population N in the environment, however the actual exposure frequency of patients was limited, that was to say, if the exposure frequency was a constant K, then

$$
\frac{dS}{dt} = -\beta \frac{S}{N} I,
$$

the equations were as follows

$$
\begin{cases}
\dfrac{dS}{dt} = -\beta \dfrac{S}{N} I \\[2mm]
\dfrac{dE}{dt} = \beta \dfrac{S}{N} I - \omega E \\[2mm]
\dfrac{dI}{dt} = \omega E - \gamma I \\[2mm]
\dfrac{dR}{dt} = \gamma I
\end{cases},
$$

and the actual incidence was standard incidence.

Referring to the design approach of Lipsitch et al,[10] this study was modeled in a population (N) of 1 000 000 individuals, consistent with the size of Wuhan city, which was approximately 9 060 000 by the end of 2019. The model simulation was based on the hypotheses: (a) the whole population was susceptible; (b) unprotected contact between the susceptible and the infectious would lead to infection; (c) the susceptible and the infectious were homogenous in the population; (d) the epidemic originated from one single case; (e) no intervention was taken in the early stage of COVID-19 epidemic; (f) the patient would be quarantined in the hospital once confirmed; and (g) deterministic dynamic model was employed. In this way, past or future dynamics of disease progression and transmission could be simulated.

3.3 | Sensitivity analysis

Sensitivity analysis is one of the commonly used methods to analyze uncertainty in the assessment. The sensitivity factors which had an important influence on the results were identified one by one from a number of uncertainty factors, and their influence degree and sensitivity degree on the overall results were analyzed and measured. The parameter whose slight variation could lead to a large change in the result was regarded as a sensitivity factor; otherwise, a nonsensitivity factor.[16] To explore the various influences of different parameters on the amount of infected people, sensitivity analysis was conducted.

The equation of sensitivity coefficient (E) was as follows

$$
E = \frac{\Delta A}{\Delta F},
$$

where ΔF denoted the percentage of uncertainty change, where the percentage of different parameters change was selected. ΔA denoted the percentage of index change, where the percentage of the peak value change of S was selected.

3.4 | Dynamic trend of effective reproductive number (R)

The basic reproductive number R_0 determines the potential of an epidemic, the extent of transmission in the absence of control measures, and the ability of control measures to decrease spread. In contrast to R_0, the effective reproductive number R measures the number of secondary cases generated by an infectious case once an epidemic is underway. $R = R_0 x$, where x is the proportion of the population susceptible,[10] which can be regarded as a constant if the whole population is susceptible. When control measures are implemented, R may decline due to the depletion of the susceptible or the containment of measures on the epidemic. In this modeling study, we calculated the R_0 of COVID-19 on the basis of reported cases from 21 January 2020 to 29 February 2020 and drew a trending chart of R.

The calculation formula of R_0[17] based on the SEIR model was

$$
R_0 = \left(1 + \frac{\lambda}{\gamma_1}\right)\left(1 + \frac{\lambda}{\gamma_2}\right)
$$

and $T_e = \frac{1}{\gamma_1}$, $T_i = \frac{1}{\gamma_2}$, $\rho = \frac{T_e}{T_g}$, then

$$
R_0 = 1 + \lambda T_g + \rho(1 - \rho)(\lambda T_g)^2,
$$

where T_g denoted the generation period, approximate to serial interval; and ρ denoted the ratio of incubation period to generation time.

We took the value of each parameter reasonably according to the above formula

$$
\lambda = \frac{\ln Y(t)}{t},
$$

where $Y(t)$ was the number of infections with symptoms by time t.

We assumed the number of confirmed cases was i, the number of suspected cases was s, and the ratio of suspected cases to confirmed cases was k, then

$$Y(t) = iks,$$

where the data of confirmed cases and suspected cases were from National Health Commission, the value of k was calculated to be 0.695 (41/59) according to the result of Huang et al[18], which showed 41 confirmed cases of the 59 suspected cases. t_1 was the date of the first confirmed case of the nCoV infection, which was 8 December 2019.[19] In this study, the value of T_g (mean 8.4 days) was assumed to be the same as Lipsitch findings[10] and the value for ρ (0.61) was derived from the average ρ of the result of Chan et al.[20]

3.5 | Data fitting

Data fitting is the process of fitting models to data and analyzing the accuracy of the curve,[21] which was performed by Python in this study. Parameters were initially set according to the released cases of COVID-19, theoretical values of t up to the current date were extracted for curve fitting, and then quadratic fitting was performed based on the actual number of potential diagnoses ($Y(t)$). A gradient equation was constructed based on the residual sum of squares of the quadratic fitting results to obtain the optimal parameters, which would be helpful in the analysis of the epidemic trend. The coefficient of determination (R^2) and fitting bias were adopted to evaluate the goodness of fitting results.

4 | RESULTS

4.1 | The distribution of confirmed COVID-19 cases

The geographical distributions of existing confirmed cases as of 29 February 2020 in Hubei Province, China, and worldwide is shown in Figure 1. The epidemic was spreading rapidly, with 27 700 existing confirmed cases in Wuhan city, 33 565 in Hubei province, 35 566 in China, and 40 903 in the whole world (Table S1).
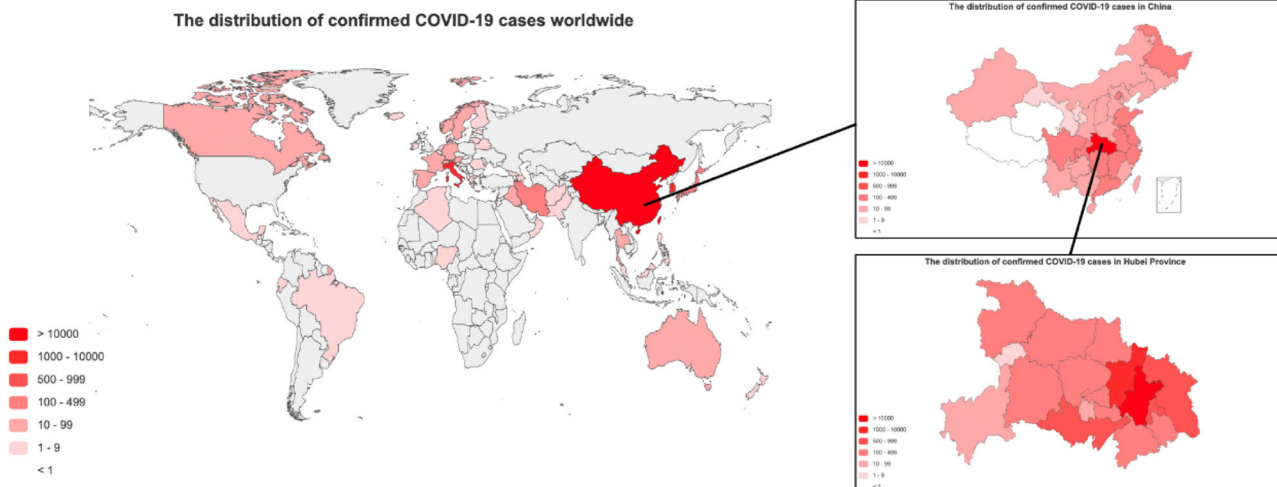
Figure 2 shows the trend of existing confirmed cases home and abroad, which indicates the peak number on 17 February 2020 in China and an obvious rise from 25 to 29 February 2020 overseas.

4.2 | Model simulation

On the basis of the previous study,[10] the case data released by Chinese authorities and official media,[22,23] the assumptions of the parameters were established, and the theoretical dynamics of COVID-19 based on the SEIR model were simulated (Figure 3).

$$N = 10000000,$$

$$\beta_0 = 0.1,$$

$$k = 10,$$

$$T_e = 7,$$

$$T_i = 10.25.$$

In the face of the outbreak, four-phase stringent measures were taken by the Chinese government (Table 1). To simulate the impact of different government control measures on the number of the infected I and the susceptible S, we assumed varied frequency of exposure (k) (Table 1) with the other parameters staying still on the four stages.

The S value curve of the susceptible population and I value curve of the infected population could be simulated as follows (Figures 4 and 5).

It could be clearly seen that the greater the government control policy, the smaller the k value, the slower the S value reduction slope of the susceptible population, and the lower the I value distribution peak of the infected population.

FIGURE 1 The distribution of existing confirmed COVID-19 cases in Hubei, China, and worldwide. COVID-19, coronavirus disease 2019
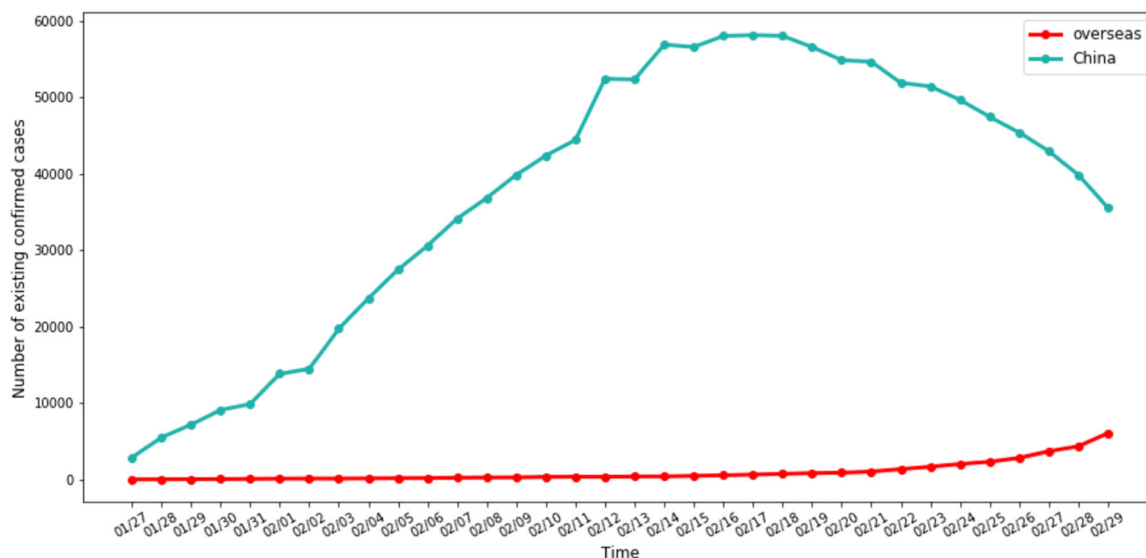
FIGURE 2 The trend of existing confirmed cases in China and overseas

The data of Figure 6 showed that there was a substantial number of new confirmed cases on 12 February and a subsequent rapid fall from 12 to 14 February, due to the addition of new diagnosis method—clinically diagnosis in Hubei province in the fourth stage of government control measures. The daily number of new recovered cases exceeded new confirmed cases in China for 12 straight days (18-29 February 2020), indicating that positive results had been achieved with the rigorous measures, so there was a subsequent downgrading of public health emergency response from the highest level to level 2 or 3 by several provinces (Gansu, Liaoning, Guizhou, Shanxi, Yunnan, Guangdong, etc) from 21 February onwards.[24]

4.3 | Sensitivity analysis

Sensitivity analysis of different parameters was performed to explore the influence degree of different parameters on S value of infected people (Figure 7).
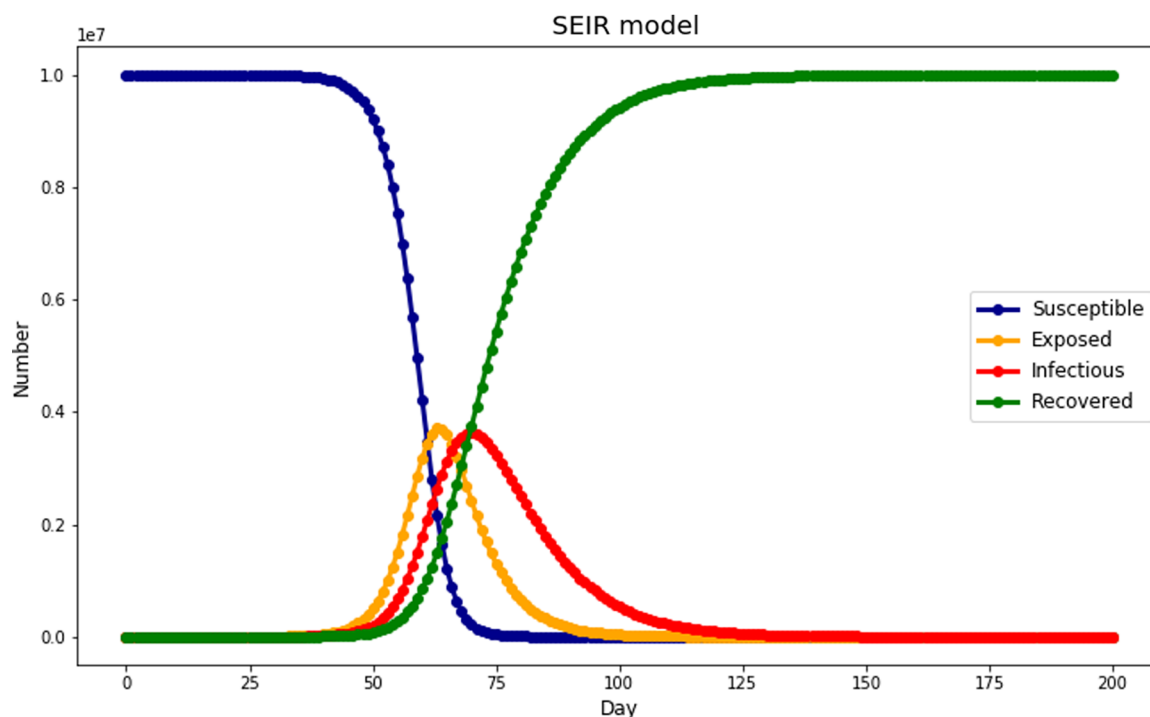

FIGURE 3 Dynamics of COVID-19 based on the SEIR model. COVID-19, coronavirus disease 2019; SEIR, susceptible-exposed-infectious-recovered
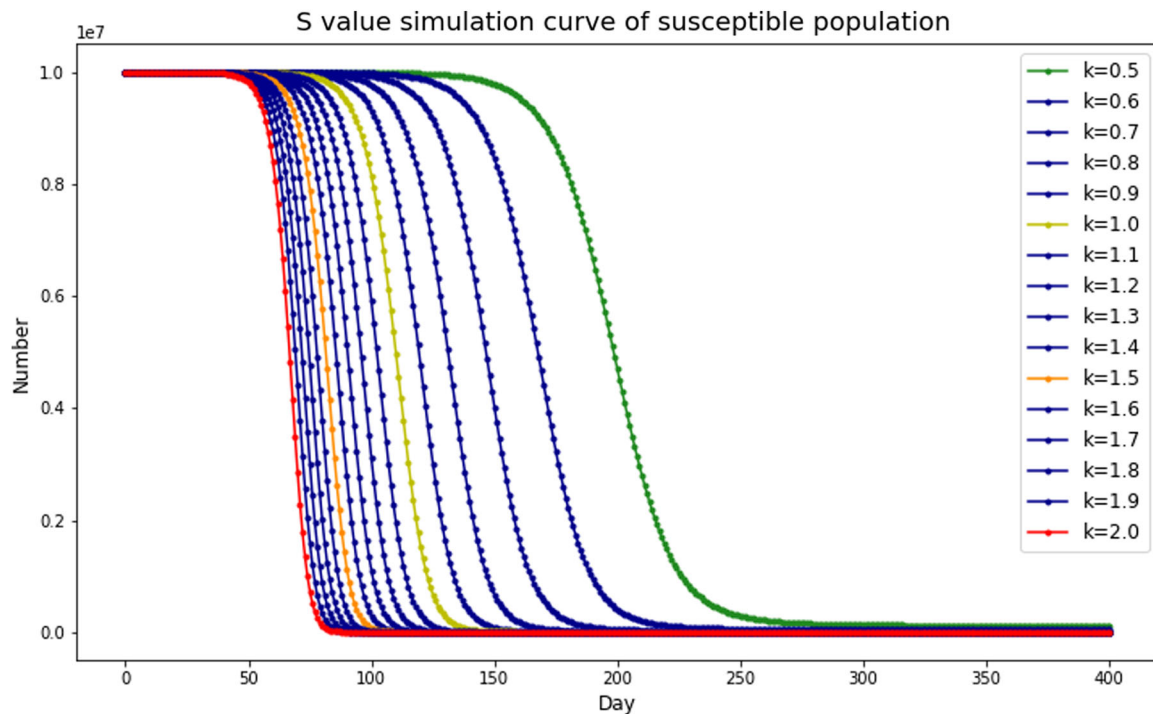
FIGURE 4 S value simulation curve of susceptible population

It could be learnt from the Figure 7 that the change of T_i value had the greatest impact on I value. The sensitivity coefficient of each parameter was $E - k = 0.4162$, $E - \beta_0 = 0.4266$, and $E - T_i = 0.8003$. Isolation and quarantine procedures would be less effective as more cases accrue, so the optimization of treatment plan and the development of specific drugs would be of more importance than isolation and protection measures.

4.4 | Dynamic trend of effective reproductive number (R)

The R values in Figure 8 were calculated on the basis of case information available from 20 January to 29 February 2020. To show a more intuitive and clear picture of R, Figure 9 was drawn to visualize the trend of R values overtime, which indicated that there was an upward trend of R from 20 January to 30 January, a downward trend from 31 January to 11 February, a short rebound on 12 February and another continuous decline from 13 to 29 February. Time-series analysis was adopted to verify the decline statistically and the result by Dickey-Fuller Test was shown in Figure 10, which proved to be nonstationary data ($P > .05$, and the data were decomposed into time series, as shown in Figure 11, whose result (No. 01 trend of Figure 11) was consistent with that of Figure 9. The feature of high infectiousness for SARS-COV-2 led to the upward trend, and No. 3 and No. 4 government measures contributed to the downward trend, the short rebound, and the subsequent decline.

4.5 | Data fitting

Data fitting (Figure 12) was conducted by polyregression using the following fitting equation

$$y = -4344.6538 + 1693.3292x + 130.7814x^2 - 3.7881x^3,$$

Mean Absolute Error (MAE): 2416.700019131622,

Mean Square Error (MSE): 8068138.399370178,

Mean Absolute Deviation (MAD): 2491.175807186635,

$R^2 = .9825472319879797$.

Then, the parameters of SEIR were adjusted by means of grid search. The results of grid search No. 1 were as follows

$$k = \text{range } (0, 10, 0.1), \beta_0 = \text{range } (0, 1, 0.01), T_i$$
$$= \text{range } (5, 20, 0.1).$$

Minimum MSE occurred when $k = 7.1$, $\beta_0 = 0.09$, and $T_i = 17.9$ (Figure 13).

The top five parameters were identified in turn (Figure 14).

Subsequently, grid search No. 2 was conducted following the process

$$k = 7.1, 0.9, 2.0, 4.0, 8.0,$$

$$\beta_0 = 0.09, 0.71, 0.32, 0.16, 0.08,$$

$$T_i = \text{range } (17.8, 18.0, 0.01).$$

TABLE 1 Different government control measures and corresponding k values

No.	Date	Government Measures	k
1	29 December 2019 to 22 January 2020	Early detection of the SARS-CoV-2 Preliminary control	$k = 2.0$
2	23 January-29 January 2020	(1) Public health level 1 response of 31 provinces (2) Strict exit screening (3) Medical support from other regions of China (4) Cancellation of mass gatherings (5) Methodological improvement on the diagnosis and treatment strategy	$k = 1.5$
3	30 January-11 February 2020	(1) Public health level 1 response of 31 provinces (2) Strict exit screening (3) Domestic and international medical support (4) The larger scale of cancellation of mass gatherings (5) Further methodological improvement on the diagnosis and treatment strategy (6) Spontaneous household quarantine by citizens (7) Two newly built hospitals' put into use (8) A clinical trial of perspective medicines	$k = 1.0$
4	12 February-20 February 2020	(1) Public health level 1 response of 31 provinces (2) Strict exit screening (3) Further medical support from home and abroad (4) Massive online teaching in a postponed semester (5) Orderly resumption of back to work (6) Addition of new diagnosis method—clinically diagnosis in Hubei province (7) Interagency mechanism (8) Further exploration of effective therapeutic strategy	$k = 0.5$

Abbreviation: SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

The final parameter was determined as

$$k = 7.1,$$

$$\beta_0 = 0.09,$$

$$T_i = 17.8,$$

whose trend was consistent with the predict linear (Figure 15)

MAE: 2627.85462448844337,

MSE: 8800640.290697332,

MAD: 2584.4329466202796,

$$R^2 = .9809627046850399.$$

When data fitting was introduced, the SEIR model simulation of COVID-19 was presented in Figure 16, with the peak existing confirmed cases of 59 769 arriving on 15 February 2020, close to the actual number of 58 016 on 17 February 2020.[25] The fitting bias was 3.02%, less than 5%.

5 | DISCUSSION

This study was mainly a data-driven analysis, with the data coming from epidemiological results of published studies (preprint included) and current case information, programed, simulated, and fitted by Python based on the SEIR model. The SEIR differs from the SIR model in the addition of a latency period and can provide a tool for predicting the size and duration of both unconstrained and managed outbreaks—the latter in the context of interventions such as case detection, quarantine, and treatment.[26]

R^2, the coefficient of determination, is used to measure the fitting degree of the regression line to the observation value and it ranges from 0 to 1. The closer the value of R^2 is to 1, the better the fitting degree of the regression line is to the observed value. The SEIR model simulation of COVID-19 after data fitting showed the peak existing confirmed cases of 59 769 arriving on 15 February 2020. The conclusion was drawn with a good coefficient of determination (R^2), which was very close to 1, and the fitting bias was 3.02% (less than 5%), suggesting high precision of the data-fitting results.
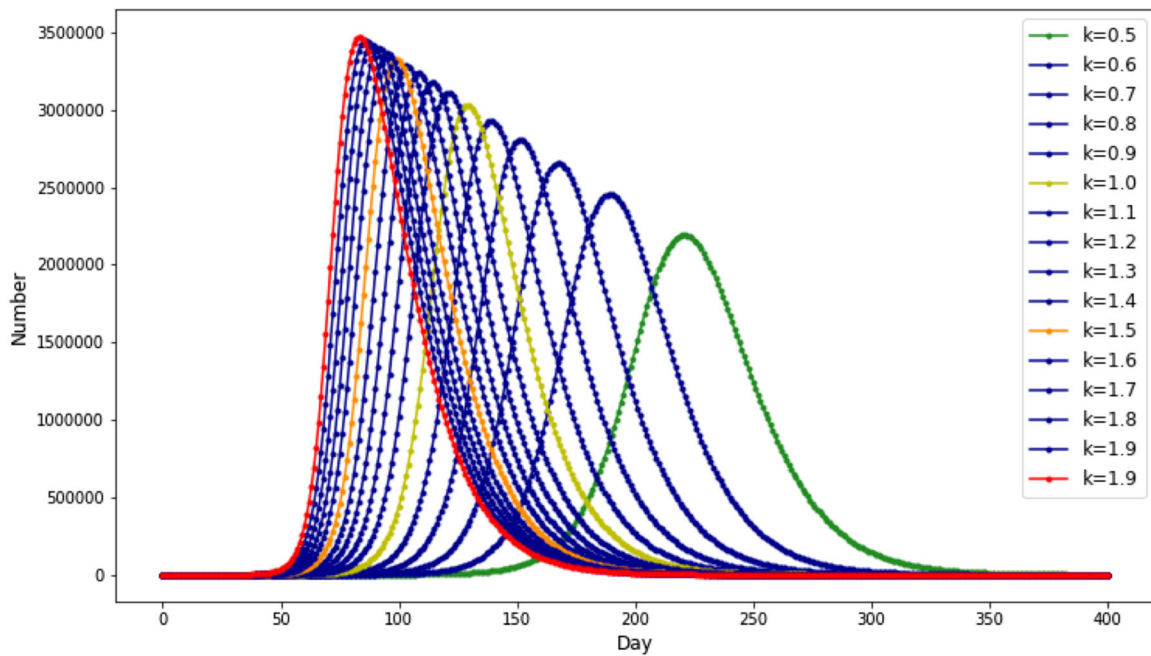
FIGURE 5 *I* value the simulation curve of the infected population

Faced with the rapidity of the growth of cases, the Chinese Health Authorities have implemented appropriate and prompt response measures, such as retrospective investigations, identifying the characteristics of the pathogen, public health level 1 response, strict exit screening, medical support (infectious disease specialists, public health professionals, physicians, nurses, surgical masks, protective suits, nucleic acid tests, etc), all round environmental sanitation and disinfection, cancellation of mass gatherings especially on the occasion of the Spring festival when it was traditional customs for people to pay new year visits, methodological improvement on the diagnosis, health communication for the public, technical guidance, clinical trial of perspective antiviral medicine like Remdesivir,[27] self-protection measures, and spontaneous household quarantine by citizens. With the presence of substantial and rigorous control measures that have
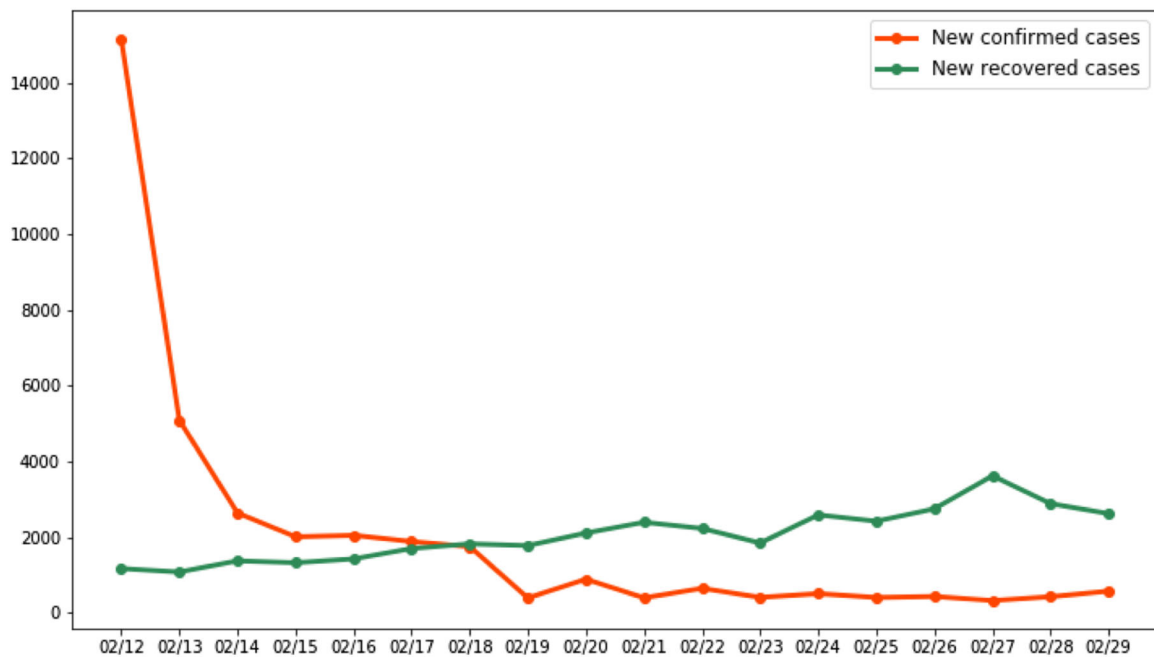

FIGURE 6 Daily number of new recovered cases and new confirmed cases in China
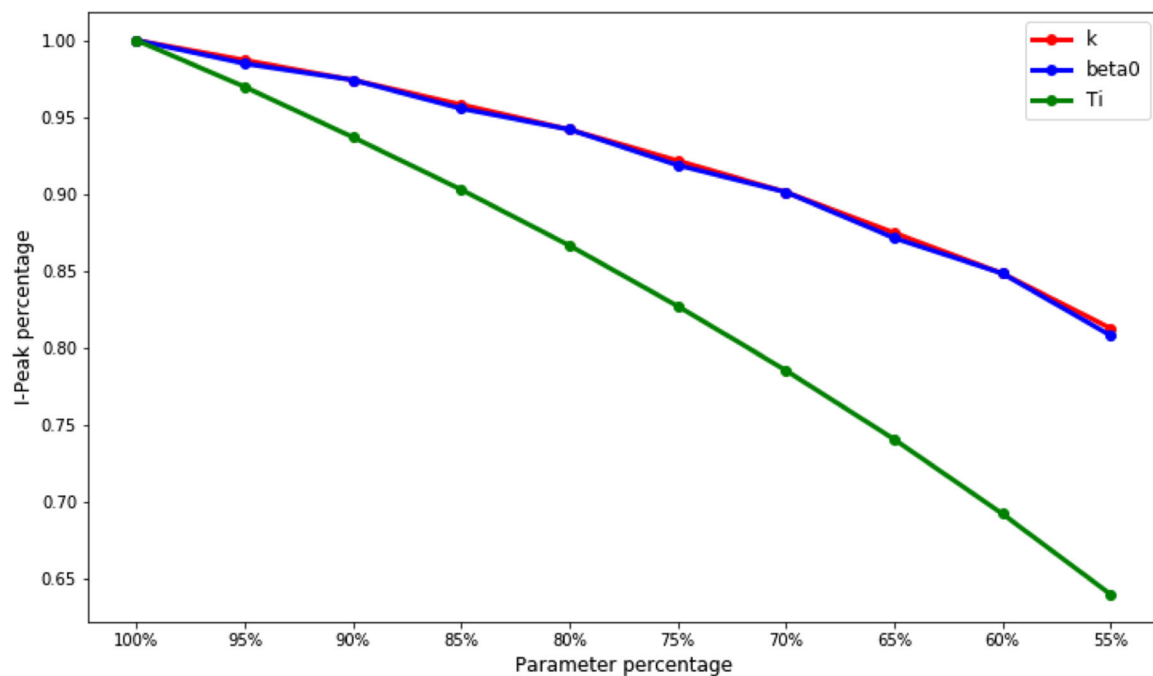
FIGURE 7 Sensitivity analysis of different parameters

been immediately carried out, we simulated the impact of different government control measures on the number of the infected I and the susceptible S by adjusting the frequency of exposure (k). The result showed that decreased frequency of exposure (K) was associated with slower rate of decline in susceptible population and slower rate of increase in infected population, indicating the possible effectiveness of intervention measures.

The parameters that could have impact on the model result contained frequency of exposure (k), probability of infection per exposure (β_0), and average recovery time (T_i). Frequency of exposure (k) could be controlled by isolation; infection per exposure (β_0) could be adjusted by wearing protective suit and average recovery time (T_i) could be shortened by the treatment strategy. To specify the most sensitive factors in intervention, we also conducted sensitivity analysis, proving that the change of T_i had stronger impact on the infected population than quarantine and protective measures, so it was urgent to explore effective treatment options. After the epidemic outbreak, Chinese scientists invested a lot of energy to carry out research to screen out clinical drugs and develop vaccines. The treatment plan at this stage focused on antiviral therapy, antibacterial therapy, glucocorticoid therapy, symptomatic treatment of fever or dyspnea, nutrition support treatment, and so on.[28-30] There was no specific antiviral strategies available, though Remdesivir, a promising drug that has been proved to be effective in fighting against Ebola virus, has been in the step of clinical trial in China.[27] Introduction of specific drug would be of great benefit to the containment of the outbreak.

It should be noted that estimation of R_0 during the pre-epidemic stage could be plagued by data uncertainty and variability,[31] so we introduced the effective reproductive number (R), whose equation

was $R = R_0x$, where x stood for the proportion of the susceptible. Since there was no vaccine at present, we assumed that everyone was susceptible. The number of infected people (including those who have been cured and quarantined) in the total population was very small, so the proportion of susceptible people could be approximately considered as constant. In this way, we established the link between R and R_0 according to the reported case data and analyzed the trend by Dickey-Fuller Test, which showed a rise from 20 to 30 January, a decline from 31 January to 11 February, a short rebound on 12 February and another continuous decline from 13 to 29 February. The rise was partly due to the onset after incubation period. The rebound was caused by inducing clinical diagnosis in the fourth phase of government measures. The number of new confirmed cases in Hubei province on 12 February was 148 40 (of which 13 332 were clinically diagnosed), which was an extraordinary large increase compared to an average of 2391 cases from 1 to 11 February (Table S2). That accounted for the temporary rebound of R. The government interventions (wearing masks, timely quarantine, improved therapeutic strategy, etc) reduced probability of infection per exposure (β_0), decreased frequency of exposure (k), and shortened the average recovery time (T_e), thus contributing to the decline of R. The declines of R could be exploited as strong evidence for the effectiveness of the interventions.

According to the data of China National Health Commission, as of 29 February, the accumulative number of cured cases was more than 41 625 nationwide, and there was a consecutive decrease of confirmed cases for 12 days. Data on 26 February showed that, for the first time, the number of new cases outside China had exceeded that of China. South Korea, Japan, Iran, and Italy were facing a surge increase in number of new confirmed cases.[32] Centers for disease

	date	confirmed	suspected	t	R	date	confirmed	suspected	t	R
0	1/20	291	54	42	2.4785x	2/10	37626.0	21675.0	63.0	2.9496x
1	1/21	437	80	43	2.56x	2/11	38800.0	16067.0	64.0	2.8997x
2	1/22	560	393	44	2.6761x	2/12	52526.0	13435.0	65.0	2.9094x
3	1/23	805	1072	45	2.8186x	2/13	55748.0	10109.0	66.0	2.8763x
4	1/24	1230	1965	46	2.9259x	2/14	56873.0	8969.0	67.0	2.8425x
5	1/25	1892	2684	47	2.986x	2/15	57416.0	8228.0	68.0	2.8089x
6	1/26	2635	5794	48	3.1055x	2/16	57934.0	7264.0	69.0	2.7758x
7	1/27	4371	6973	49	3.1476x	2/17	58016.0	6242.0	70.0	2.7425x
8	1/28	5761	9239	50	3.1748x	2/18	57805.0	5248.0	71.0	2.7093x
9	1/29	7439	12167	51	3.1973x	2/19	56303.0	4922.0	72.0	2.6747x
10	1/30	9331	15238	52	3.2075x	2/20	54965.0	5206.0	73.0	2.6429x
11	1/31	11315	17988	53	3.2043x	2/21	53284.0	5365.0	74.0	2.6107x
12	2/1	13775	19544	54	3.1902x	2/22	51606.0	4148.0	75.0	2.5764x
13	2/2	16400	21558	55	3.1762x	2/23	49824.0	3434.0	76.0	2.5436x
14	2/3	19414	23214	56	3.1595x	2/24	47672.0	2824.0	77.0	2.5105x
15	2/4	22974	23260	57	3.1362x	2/25	45604.0	2491.0	78.0	2.479x
16	2/5	26334	24702	58	3.1161x	2/26	43258.0	2358.0	79.0	2.4476x
17	2/6	29017	26359	59	3.0915x	2/27	39919.0	2308.0	80.0	2.4132x
18	2/7	31774	27657	60	3.0654x	2/28	37414.0	1418.0	81.0	2.3798x
19	2/8	33738	28942	61	3.0355x	2/29	35329.0	851.0	82.0	2.3492x
20	2/9	35982	23589	62	2.9878x					

FIGURE 8   R values from 20 January to 29 February 2020

control and prevention of United States had confirmed an instance of community spread of COVID-19 in California, without any relevant travel history or exposure to another known patient with COVID-19.[33] Given the potential of pandemic, governments worldwide are scrambling to prevent the spread of the new coronavirus. Under a strict mechanism of massive prevention and control, China has seen a rapid decrease in new cases of coronavirus. The unique measures China has taken may be referential and enlightening to other countries with rapid spread of the virus.

Evidence from the four-phase stringent measures showed that it was foremost to ensure early detection, early isolation, and early treatment, cutting off the spread from the upstream. This could be achieved by the public health emergency response and interagency mechanism. China's public health emergency response is categorized into four levels, with level I being the highest. Level 1 means the provincial disease control headquarters, in the event of a particularly serious public health emergency, will conduct emergency response work under the decisions and command of the State Council. Initiating public health emergency response and establishing a nationwide interagency mechanism was beneficial for monitoring the implementation of the early detection, early isolation, and early treatment. Another critical factor in enhancing early detection was the improvement of diagnosis method. Nucleic acid tests accounted for a big proportion of the initial pathogenic diagnosis techniques because of the convenience and rapidity, nevertheless, nucleic acid tests had returned false-negative results for people infected with the nCoV, with a 30% to 50% positive rate. Those patients with pneumonia symptoms and virus infection by the lung image but negative result by nucleic acid test were called "false-negative." Caution should be taken against "false-negative" because its harm was self-evident. It would not only delay the timely treatment of infected patients, but also cause some patients with less symptoms to
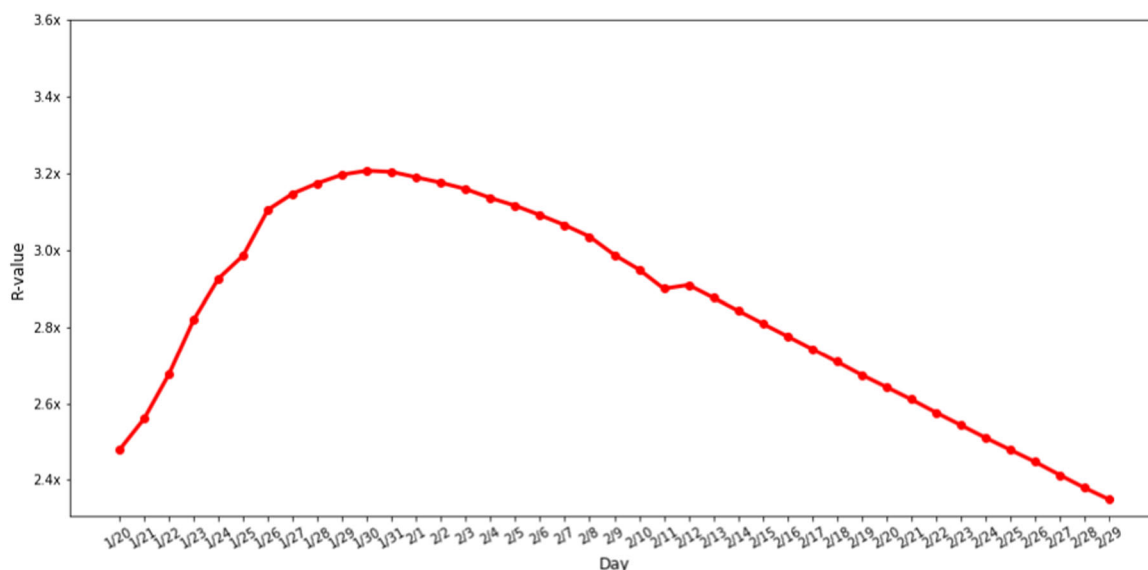
FIGURE 9 The trend of *R* values overtime

```
Results of Dickey-Fuller Test:
Test Statistic               -1.259368
p-value                       0.647550
#Lags Used                    2.000000
Number of Observations Used  38.000000
Critical value (1%)          -3.615509
Critical value (5%)          -2.941262
Critical value (10%)         -2.609200
dtype: float64
```

FIGURE 10 Results of Dickey-Fuller test

circulate in the society because they thought they were not infected so as to increase the probability of other people's infection. Therefore, it was of great importance to introduce the clinical diagnosis in addition to nucleic acid tests in the fourth phase of government measures. The Chinese health authorities issued the fifth version of diagnosis and treatment plan for COVID-19.[34] According to the updated criteria, suspected cases with imaging characteristics of pneumonia could be clinically diagnosed as COVID-19 in Hubei province. In this way, patients could receive standardized treatment as soon as possible, aiming to further improve the success rate of treatment and contain the spread of the epidemic. Second, patients
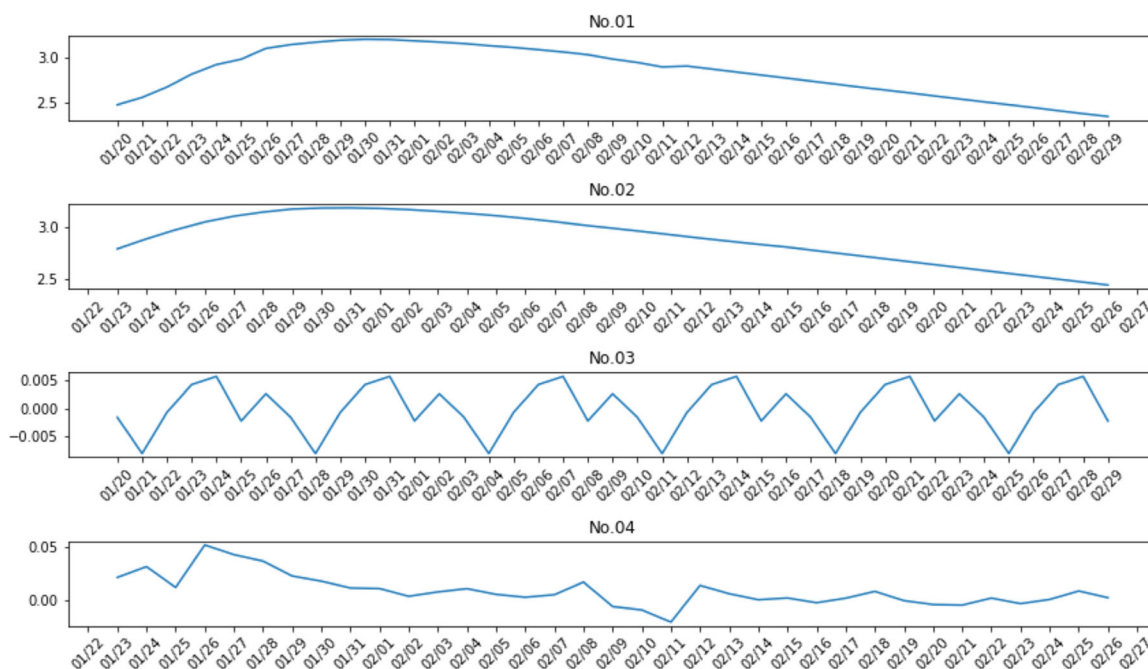

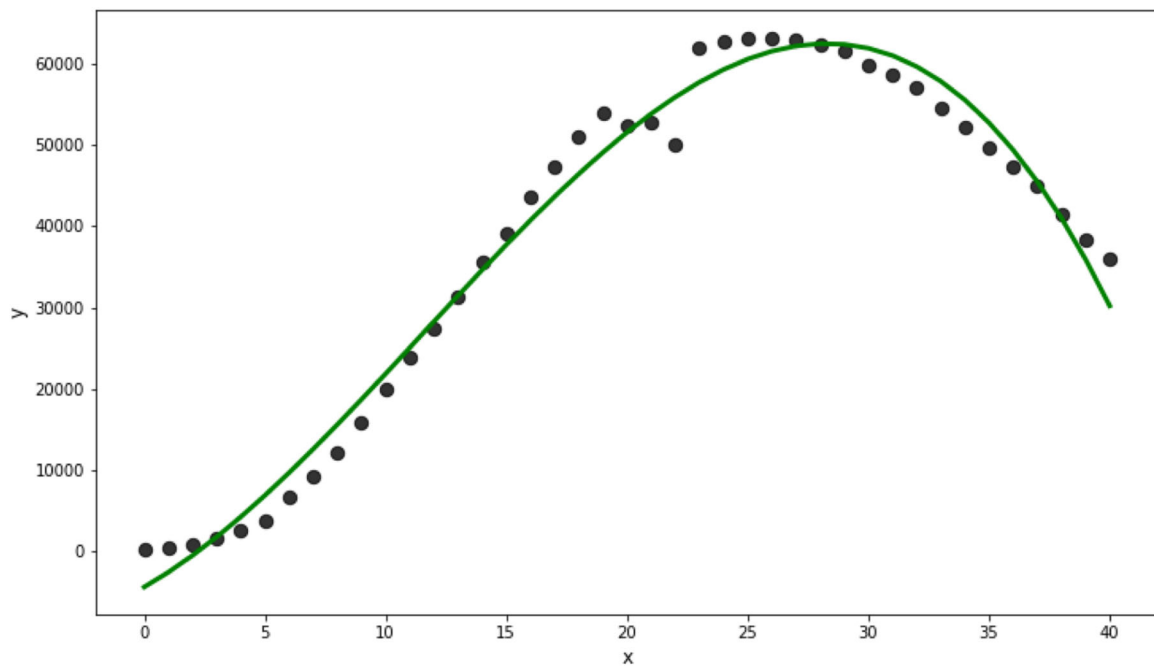
FIGURE 11 Data decomposed into time series

FIGURE 12 Data fitting by polyregression

should be admitted to designated hospitals and comprehensive treatment strategy including antiviral therapy, respiratory support, circulatory support, and immunity enhancement should be adopted for those who were critically ill. The initial surge of cases in Wuhan made it difficult for the hospital capacity to meet the demand; therefore, beside the current designated medical institutions, Wuhan learned from Beijing's successful experience during the fight against SARS in 2003 and built two new temporary hospitals—Huoshenshan Hospital and Leishenshan Hospital, which were designed to have a capacity of 1000 beds and 1500 beds, respectively. Third, adequate medical supplies, such as health professionals, masks, protective suits, and coronavirus test kits, should be guaranteed, otherwise
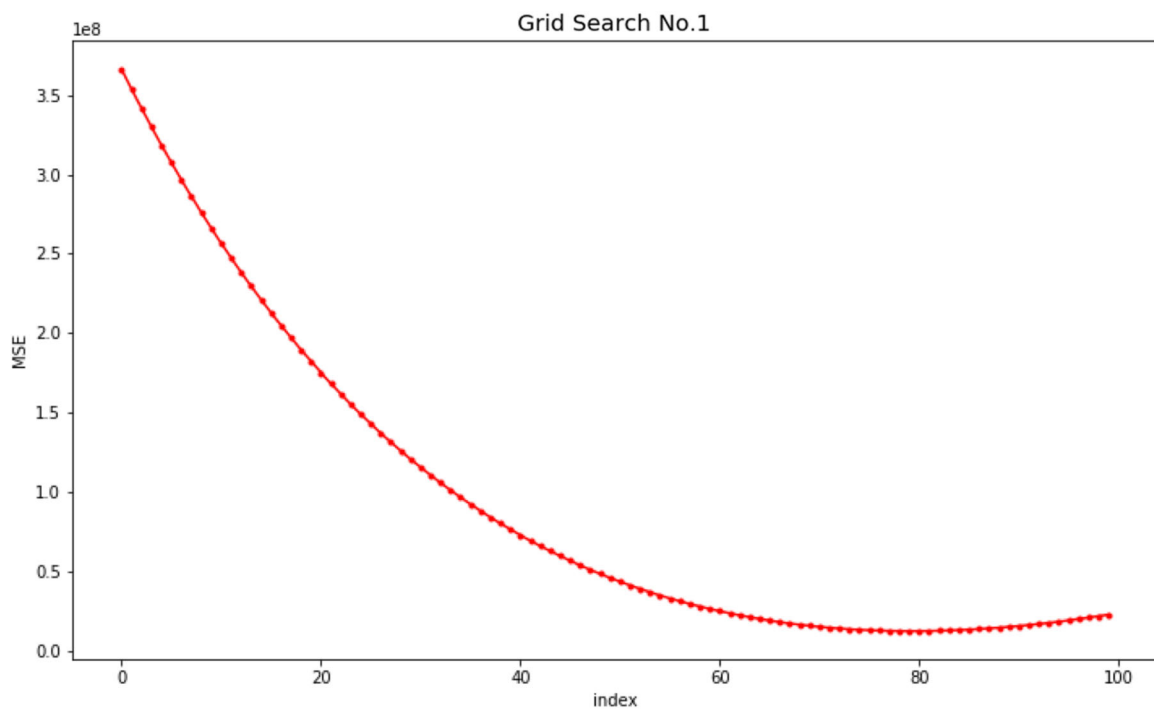

FIGURE 13 Result of grid search No. 1

FIGURE 14 The top five parameters. MSE, mean square error

	index	MSE
1066479	[7.1000000000000005, 0.09, 17.900000000000002]	1.235485e+07
145779	[0.9, 0.71, 17.900000000000002]	1.235485e+07
304929	[2.0, 0.32, 17.900000000000002]	1.236190e+07
602529	[4.0, 0.16, 17.900000000000002]	1.236190e+07
1201329	[8.0, 0.08, 17.900000000000002]	1.236190e+07

insufficient medical resources may resulted in the sharp rise of the COVID-19 patients.

Faced with the massive public health crisis, the authorities and the public are both required to join the fight against the nCoV. Continuous interventions and increased vigilance are needed to prevent the ongoing spread of the nCoV. More domestic and global collaborative efforts are required in the way to fight against the COVID-19 to reduce the threat of the outbreak.

6 | STRENGTH

Though there were several studies concentrating on the R_0 and spread of the epidemic, to the best of our knowledge, there was a little evidence about the dynamic R trend and the effectiveness of government control measures. Also, this study reflected the latest updates about the epidemic up to 29 February 2020. In contrast to a single R_0, we attempted to take into account the dynamic changes of R values on different scenarios of the epidemic. This study also

established the link between R values and the effect of government measures. Besides, we adopted sensitivity analysis to identify the most effective factor in containing the epidemic. In addition, the estimate of transmission dynamics of the COVID-19 was calculated by simulation followed by data fitting, which could be helpful in enhancing the accuracy and precision of the prediction. Finally, given the situation of pandemic potential, the four-phase stringent measures by Chinese government could be referential and enlightening to the other countries suffering from COVID-19.

7 | LIMITATION

We were still in the stage of much uncertainty about the COVID-19, both in the actual virus host and the possible outbreak scale, so the accuracy and precision of our estimate were subjected to a lot of factors; therefore, it would be beneficial for further studies to offer more valuable and precise information about the novel virus. The actual number of infected patients was not exactly the same as the
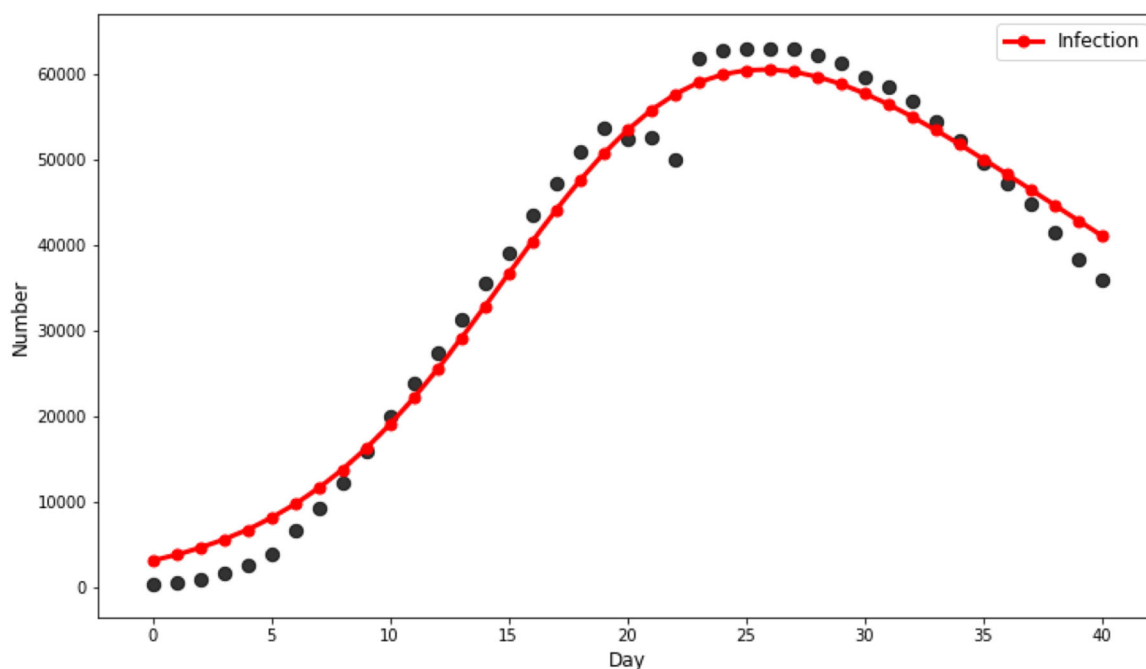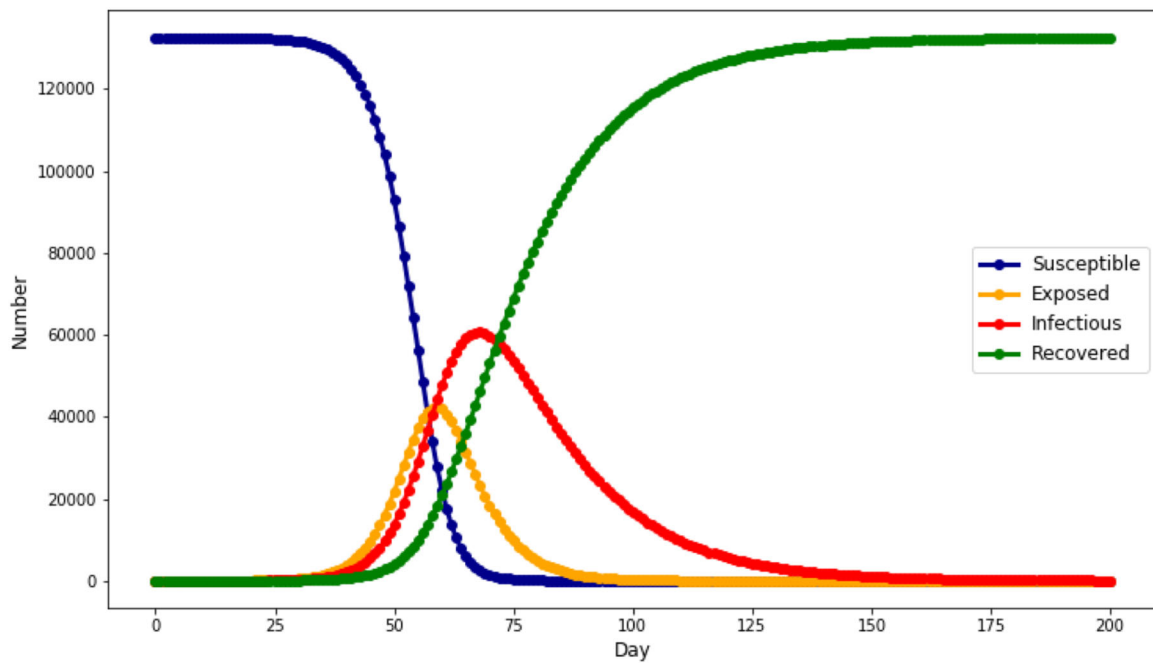

FIGURE 15 Predict linear

FIGURE 16 The SEIR model simulation of COVID-19 after data fitting. COVID-19, coronavirus disease 2019; SEIR, susceptible-exposed-infectious-recovered

number of reported cases on that day, and there may be some missing reports and undetected cases. Since the actual number of infected and suspected patients was not observable, there must be some unavoidable discrepancy between the confirmed patients and actual number of infection, which lead to the inevitable deviation of the estimation. Though we adopted the model in the addition of latency, it did not account for the infectiousness during incubation period, so it may reduce the accuracy as a consequence. The travel of people flow between different regions was not well considered, especially the migration after the Spring Festival—back to work and back to school. Sensitivity analysis suggested that effective drug therapy could have great influence on the intervention result, it would have been more comprehensive if the study could have included the effect of antiviral treatment. Finally, our research was a data-driven analysis, with data coming from heterogeneous sources, which would definitely introduce inevitable bias, therefore care should be taken in the interpretation of the results.

8 | CONCLUSION

By simulation and data fitting, the model showed the peak existing confirmed cases of 59 769 arriving on 15 February 2020, with sufficient goodness of fit. The optimization of therapeutic strategy and the development of specific drugs would be of more importance than quarantine and protective procedures as more cases accrue. The declines in the dynamic trend of effective reproductive number (*R*) indicated the effectiveness of four-phase stringent government measures. It is an unprecedented outbreak of coronavirus in the 20th century, which is not exactly the epidemic of a single country, more a public concern of the whole world; therefore, more international cooperation are required to combat the nCoV, focusing on both persistent strict domestic interventions and vigilance against exogenous imported cases.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Yaqing Fang http://orcid.org/0000-0002-9963-4462

REFERENCES

1. Li Q The 2019-nCoV Outbreak Joint Field Epidemiology Investigation Team. An Outbreak of NCIP (2019-nCoV) Infection in China—Wuhan, Hubei Province, 2019-2020. *China CDC Weekly*. 2020;2(5):79-80.
2. Tan WJ, Zhao X, Ma XJ, et al. A novel coronavirus genome identified in a cluster of pneumonia cases—Wuhan, China 2019–2020. *China CDC Wkly*. 2020;2(4):61-62. http://weekly.chinacdc.cn/en/article/ccdcw/2020/4/61
3. World Health Organization. WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020. Accessed February 11, 2020.
4. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020. https://doi.org/10.1038/s41586-020-2012-7
5. World Health Organization. Statement on the Second Meeting of the International Health Regulations. https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov). Accessed February 4, 2020.

6. National Health Commission of the People's Republic of China. The updates of COVID-19 as of 29 February, 2020. http://www.nhc.gov.cn/xcs/yqtb/202003/9d462194284840ad96ce75eb8e4c8039.shtml. Accessed March 1, 2020.

7. World Health Organization. Coronavirus disease 2019 (COVID-19) situation report (40). https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200229-sitrep-40-covid-19.pdf?sfvrsn=849d0665_2. Accessed March 1, 2020.

8. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. [published online ahead of print January 29, 2020]. N Engl J Med. https://doi.org/10.1056/NEJMoa2001316

9. Roberts MG, Nishiura H. Nishiura, early estimation of the reproduction number in the presence of imported cases: pandemic influenza H1N1-2009 in New Zealand. PLoS One. 2011;6:e17835.

10. Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science. 2003; 300(5627):1966-1970.

11. Read JM, Bridgen JRE, Cummings DAT, Ho A, Jewell CP. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. medRxiv. 2020. https://doi.org/10.1101/2020.01.23.20018549. Preprint online January 24, 2020.

12. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet. 2020;395:689-697.

13. Nature. Coronavirus latest: scientists scramble to study virus samples. https://www.nature.com/articles/d41586-020-00154-w. Accessed February 4, 2020.

14. Zhao S, Lin Q, Ran J, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. Int J Infect Dis. 2020;92:214-217.

15. Ma ZE, Zhou YC, Wang WD, et al. Mathematical Modeling of Epidemic Dynamics. Beijing, China: Science Press; 2004.

16. Qian G, Mahdi A. Sensitivity analysis methods in the biomedical sciences. Math Biosci. 2020. https://doi.org/10.1016/j.mbs.2020.108306

17. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc R Soc B. 2007;274:599-604.

18. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan,China. Lancet. 2020;395(10223): 497-506. https://doi.org/10.1016/S0140-6736(20)30183-5

19. Chinanews. The progress of COVID-19 by Wuhan Health Commission. http://www.chinanews.com/shipin/2019/12-31/news843209.shtml. Accessed February 5, 2020.

20. Chan JF, Yuan S, Kok KH, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020;395:514-523.

21. Xie Z. Data fitting and scenario analysis of vaccination in the 2014 Ebola outbreak in Liberia. Osong Public Health Res Perspect. 2019; 10(3):187-201.

22. Chinanews. National Health Commssion: the average latency of 2019-nCoV is about 7 days. http://www.chinanews.com/gn/2020/01-22/9067014.shtml. Accessed February 2, 2020.

23. Sina. Beijing has discharged nine patients with pneumonia infected with the new coronavirus, with an average stay of 10.25 days. http://news.sina.com.cn/c/2020-02-01/doc-iimxxste8097489.shtml. Accessed February 3, 2020.

24. Chinadaily. Several provinces have downgraded public health emergency response level. http://cnews.chinadaily.com.cn/a/202002/24/WS5e531de4a3107bb6b57a1bef.html. Accessed February 28, 2020.

25. National Health Commission of the People's Republic of China. The updates of COVID-19 as of 17 February, 2020. http://www.nhc.gov.cn/xcs/yqtb/20200/261f72a74be14c4db6e1b582133cf4b7.shtml. Accessed March 3, 2020.

26. Getz WM, Salter R, Mgbara W. Adequacy of SEIR models when epidemics have spatial structure: Ebola in Sierra Leone. Philos Trans R Soc Lond B Biol Sci. 2019;374(1775):20180282.

27. Ministry of Science and Technology of the People's Republic of China. Press Conference of 2019-nCoV by National Health Commission of People's Republic of China. http://www.most.gov.cn/kjbgz/202002/t20200205_151401.htm. Accessed February 4, 2020.

28. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 Novel coronavirus-infected pneumonia in Wuhan, China. JAMA. 2020. https://doi.org/10.1001/jama.2020.1585

29. Jin YH, Cai L, Cheng ZS, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). Mil Med Res. 2020; 7(1):4.

30. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395: 497-506.

31. Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the basic reproduction number (R_0). Emerging Infect Dis. 2019;25: 1-4.

32. Hindustantimes. More novel coronavirus cases reported daily outside China than inside: WHO. https://www.hindustantimes.com/world-news/coronavirus-outbreak-update-more-novel-coronavirus-cases-reported-daily-outside-china-than-inside-who/story-VN2iFxRjbD2C7rAIr1eVUP.html. Accessed February 28, 2020.

33. Centers for disease control and prevention. CDC confirms possible instance of community spread of COVID-19 in U.S. https://www.cdc.gov/media/releases/2020/s0226-Covid-19-spread.html. Accessed February 28, 2020.

34. Government of People's Republic of China. The diagnosis and treatment plan for NCIP. http://www.gov.cn/zhengce/zhengceku/2020-02/05/content_5474791.htm. Accessed February 5, 2020.

How to cite this article: Fang Y, Nie Y, Penny M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. J Med Virol. 2020;92:645–659. https://doi.org/10.1002/jmv.25750