Coronavirus Pandemic

Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis

Rui Huang[1], Miao Liu[2#], Yongmei Ding[1#]

[1] *Department of Mathematics and Statistics, College of Science, Wuhan University of Science and Technology, Wuhan, Hubei Province, China*
[2] *Department of Pathology, Brigham and Women's Hospital, Harvard Medical School Boston, MA, United States*

\# These authors contributed equally.

Abstract

Currently, the outbreak of COVID-19 is rapidly spreading especially in Wuhan city, and threatens 14 million people in central China. In the present study we applied the Moran index, a strong statistical tool, to the spatial panel to show that COVID-19 infection is spatially dependent and mainly spread from Hubei Province in Central China to neighbouring areas. Logistic model was employed according to the trend of available data, which shows the difference between Hubei Province and outside of it. We also calculated the reproduction number R_0 for the range of [2.23, 2.51] via SEIR model. The measures to reduce or prevent the virus spread should be implemented, and we expect our data-driven modeling analysis providing some insights to identify and prepare for the future virus control.

Introduction

The new type of pneumonia, which started in Wuhan, Hubei Province, China in December 2019, has infected 114,350 cases and killed 4,023 in 107 countries/regions (by 12:56pm on March 10, 2020) [1]. The virus causing the pneumonia has been named new coronavirus 2019 (2019-ncov), and renamed as COVID-19 based on a tweet by the World Health Organization (WHO) sent on February 11 [2]. Being highly contagious, it leads to tens of thousands of cases and caused great panic worldwide. Even if some control measures have been taken across China, such as home quarantine, wearing masks, and so on, the number of confirmed cases still showed a continuous increasing trend. Much remains unclear and to be investigated at this stage.

While many doctors and nurses were fighting the uphill battle in the frontline, scientific researchers from various fields also played an important role in conducting research on Virus prevention, transmission mechanisms and drug treatment and development [3-11]. Rapid spread also caused a heavy psychological impact on the public [12], 98.54% of respondents expressed great fear due to high contagiousness (64.71%) and lack of effective treatment (19.92%) [13]. Mathematical and statistical analysis are also employed for containment strategies, representing a great challenge to the modelers as huge numbers of infected patients caused a lack of clinical resources and limited data, especially at the center of the infectious district [14,15].

In this report we updated the data sequence and provided an analysis of real-time data from the websites of People's Daily, CCTV news, Hubei Daily and Lilac Garden to describe the spatial correlation, transmission model and dynamics characteristics. We applied the Moran index a strong statistical tool to a spatial panel, and we found that COVID-19 infections are spatially dependent and mainly spread from Hubei Province in Central China to neighboring areas. The cumulative curves of infection cases are employed for more stable and reliable prediction via Logistic model as the short time-series is affected by irregularities and reporting lags. Two or more secondary cases generated by an index case under the condition of not considering the "super spreaders", which urges the Chinese people to continue to work hard to control the virus and reduce

the basic regeneration number to below 1 as soon as possible.

Model, methods and data

Spatial dependent structure

Attribute values of adjacent region units are often characterized by Moran index. Suppose the variable x is the attribute value of the region, the global Moran index "I" can be expressed as (1)

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \sum_{k=1}^{n}(x_k - \bar{x})^2} = \frac{\sum_{i=1}^{n} \sum_{j \neq i}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{s^2 \sum_{i=1}^{n} \sum_{j \neq i}^{n} w_{ij}} \quad (1)$$

Where $s^2 = \frac{1}{n}\sum_i (x_i - \bar{x})^2$, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, and w_{ij} is the weight matrix of spatial dependence to describe the correlation between the explained variables of section i and section j. It is evident that the range of Moran index is [-1, 1], and value zero means no correlation between two parts. There are two general descriptions for the spatial weight matrix: one is the matrix based on geographical distance, where w_{ij} means the physical distance between i and j district; another one is a simple binary adjacency matrix, whose elements in row i and column j can be expressed as:

$$W_{ij} = \begin{cases} 1 \ region \ i \ and \ j \ are \ connected \\ \qquad\quad 0 \ others \end{cases} \quad (2)$$

Logistic regression model

Logistic function or Logistic curve is a kind of s-shaped function, which is namely an ordinary differential equation as shown in (3).

$$\frac{dP}{dt} = rP(1 - \frac{P}{k}) \quad (3)$$

where P means the probability of personal infection; r means the increasing rate of cases, which measures the speed of curve change, and t means the time span. The solution of above differential equation is shown in (4).

$$P(t) = \frac{KP_0 e^{rt}}{K + P_0(e^{rt} - 1)} \quad (4)$$

where P_0 is the initial value of P, K is the maximum capacity of the environment.

SEIR model

The SEIR model has been used to predict the virus's rate of spread since the new pneumonia outbreak at the end of 2019. If an individual in an infected state (I) contacts an individual in a susceptible state (S), then the probability of the susceptible individual being infected

is β; Individuals in the incubation period (E) become infected with probability γ_1 in unit time (day) (I); An individual in an infected state (I) is converted to a cured state (R) with probability γ_2 per unit time (day). Then the transmission process of COVID-19 can be described by the following differential equations:

$$\begin{cases} \frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N} \\ \frac{dE(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma_1 E(t) \\ \frac{dI(t)}{dt} = \gamma_1 E(t) - \gamma_2 I(t) \\ \frac{dR(t)}{dt} = \gamma_2 I(t) \end{cases} \quad (5)$$

where S(t), E(t), I(t) and R(t) represent the number of individuals in susceptible, latent, infected and recovered states at time t respectively, and N represents the total number of individuals, N = S(t)+E(t)+I(t)+R(t). The basic regeneration number can be expressed as:

$$R_0 = \left(1 + \frac{\lambda}{\gamma_1}\right) + \left(1 + \frac{\lambda}{\gamma_2}\right) \quad (6)$$

where $\lambda = \frac{\ln(Y_t)}{t}$ refers to the increasing rate during early exponential growth, Y_t is the number of infections with symptoms as of time t. Incubation and infection period is expressed as $T_E = \frac{1}{\gamma_1}$ and $T_I = \frac{1}{\gamma_2}$, Generation time is approximated as sequence interval, *i.e* $T_g = T_E + T_I$. If $\rho = \frac{T_E}{T_g}$ is the proportion of incubation period to generation time, the basic regeneration number can be expressed as:

$$R_0 = (1 + \lambda T_g) + \rho(1 - \rho)(\lambda T_g)^2 \quad (7)$$

About data

Data on COVID-19 were collected in China between January 13, 2020 to March 9, 2020. All real-time dynamic data were extracted from the open website of People's Daily, CCTV news, Hubei daily and Lilac Garden.

Real data analysis

Spatial feature

We provide the heat map of the regional distribution of COVID-19 as of March 9 in China, which is shown in Figure 1. Different colors are borrowed for showing the number of confirmed cases in different provinces. The darker the color, the more severe the infection. The disease has obvious regional characteristics in terms of regional distribution among 31 provinces and cities in China, with significant spatial agglomeration characteristics.

It's worth noting that the infection areas are mainly distributed in Hubei province and spread to neighboring provinces. Wuhan, the capital of Hubei province and also the worst-hit city (Figure 2), had 49,965 infected cases as of March 9. Deep color shows more infectious cases. In order to control the spread of the epidemic, Chinese government has taken measures to cut off the source of infection, such as lockdown the city of Wuhan, isolate measures for people returning from Wuhan or who have come into contact with Wuhan people.

In order to test the spatial autocorrelation of COVID-19 in China, the interaction between provinces and cities has been estimated through Moran index from January 15, 2020 to January 31, 2020 by using the formula (1). Results are shown in Table 1.

Moran coefficients based on adjacent weight matrix are all significant at the significance level of 1% (the third column in Table 1), and the values are around the interval of [0, 1], since there is a positive correlation among the confirmed cases according to the geographical structure, and its spatial distribution has obvious agglomeration characteristics (Figure 3). The increase of confirmed pneumonia cases in one region will inevitably lead to the increasing cases in adjacent regions, which means that a positive spillover effect occurs.

Comparatively, no significant spatial correlation is tested out based on spatial geographic distance (the last column of Table 1), which indicates the spreading direction in China is mainly based on adjacent areas to neighbors, and doesn't matter how far the distance to the

Figure 1. China epidemic of COVID-19.
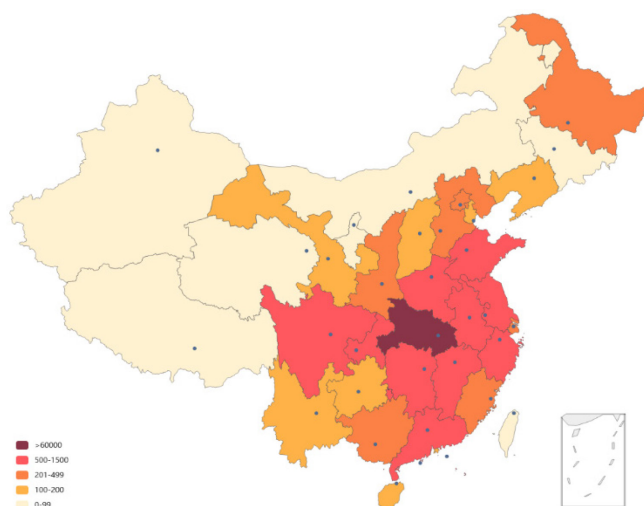
The heat map of the regional distribution of COVID-19 as of March 9 in China. The darker the color, the more severe the infection.

infectious centre. So provinces adjacent to Hubei are at higher risk, the same as the cities adjacent to Wuhan.

In Figure 3, the confirmed cases are expressed by the abscissa and its spatial lag vector explained by the ordinate, and it shows the positive correlation about the infectious cases. The first quadrant is the (high, high) space, which indicates that the provinces and cities with high number of confirmed pneumonia cases are surrounded by the same cases' situation of provinces and cities; The third quadrant is a low spatial association, presents that provinces and cities with low

Table 1. Moran Index and its test.

Date	Measurement by the adjacency		Measurement by geographical distance	
	Moran'I	P-value	Moran'I	P-value
31	0.0268	0.0001	-0.0318	0.4403
30	0.0291	0.0002	-0.0302	0.4057
29	0.0488	0.0007	-0.0256	0.3161
28	0.0466	0.0016	-0.0265	0.3351
27	0.0491	0.0017	-0.0296	0.3969
26	0.0395	0.0016	-0.0233	0.2844
25	0.0252	0.0022	-0.0261	0.3382
24	0.0184	0.0006	-0.0323	0.4543
23	0.0394	0.0016	-0.0428	0.6724
22	0.0252	0.0022	-0.0452	0.7267
21	0.0184	0.0006	-0.0490	0.7948
20	0.0139	0.0010	-0.0503	0.8198
19	0.0088	0.0006	-0.0507	0.8324
18	$1.999*10^{-3}$	$2.610*10^{-5}$	-0.0499	0.8205
17	$2.813*10^{-3}$	$9.284*10^{-6}$	-0.0499	0.8211
16	$2.091*10^{-3}$	$4.020*10^{-6}$	-0.0499	0.8217
15	$5.419*10^{-4}$	$2.704*10^{-6}$	-0.0499	0.8217

**Figure 2.** The epidemic in Hubei province, March 9, 2020.
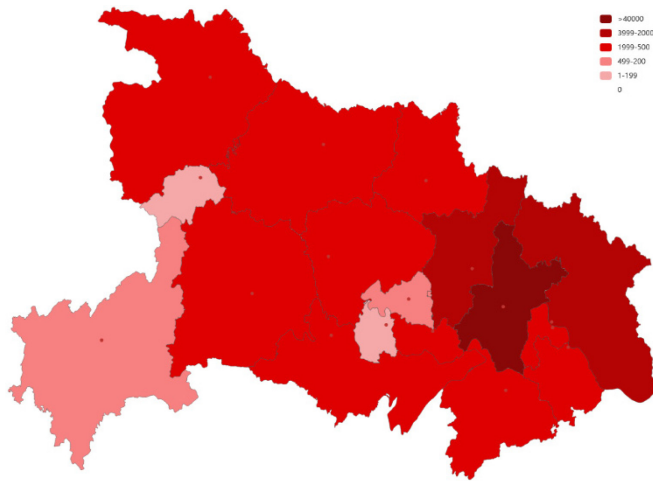
pneumonia diagnoses are surrounded by the same situation of provinces and cities, and an obvious spatial agglomeration characteristic is captured, with a strong spatial dependence. It is consistent with the results of Moran I test correspondingly (column 2 in Table 1).

*Epidemic prediction based on Logistic model*

Although COVID-19 spread rapidly, we recognized that the transmission process takes place in a limited population. The transmission rate will decline after reaching a certain threshold with the implementation of various containments. This process is completely consistent with the Logistic curve. In the following analysis, we divided the data into two stages according to the different months. The first stage is from the date

Figure 4. Real cases, fitting curves and prediction for the confirmed cases (January).
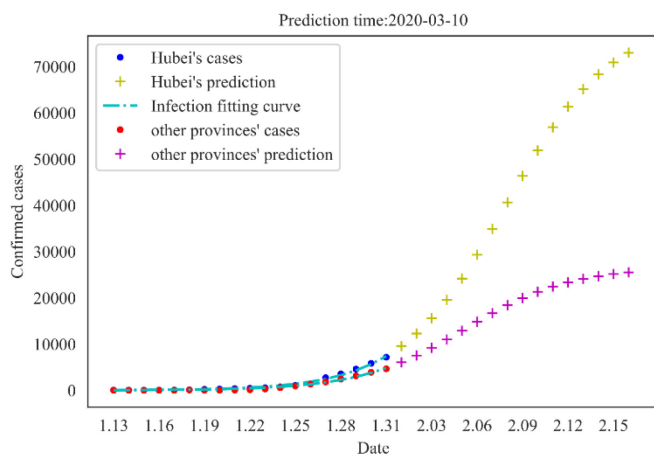
The test set node was from the date of 13 to 31 of January. Logistic model was employed and the average error was -1.6%.

Figure 3. Regional correlation diagnosis of ongoing COVID-19 in China.
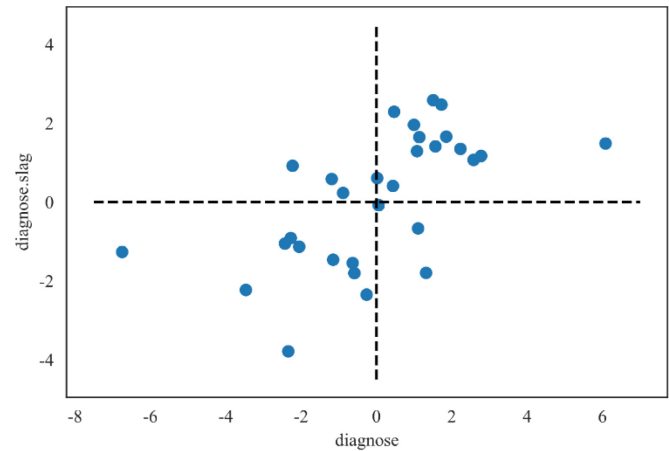
The first quadrant is the (high, high) space, which indicates that the provinces and cities with high number of confirmed pneumonia are surrounded by the same situation; The third quadrant is a low spatial association, presents that provinces and cities with low pneumonia diagnoses are surrounded by the same situation.

of 13 to 31 in January. A logistic model was employed and the average error was -1.6%, which shows reliability of our model. Confirmed cases are chosen to compare between other provinces and Hubei, and figure out that they coincided at the beginning of January, but they were increased in Hubei province at the end of the same month. This trend is due to the fact that medicine resources were lacking and the data were limited to the early stage of the infection in this province. Later, medical resources from other provinces all supported Hubei province for more efficient diagnosis leading to

Figure 5. Real cases, fitting curves and prediction for the confirmed cases (February-).
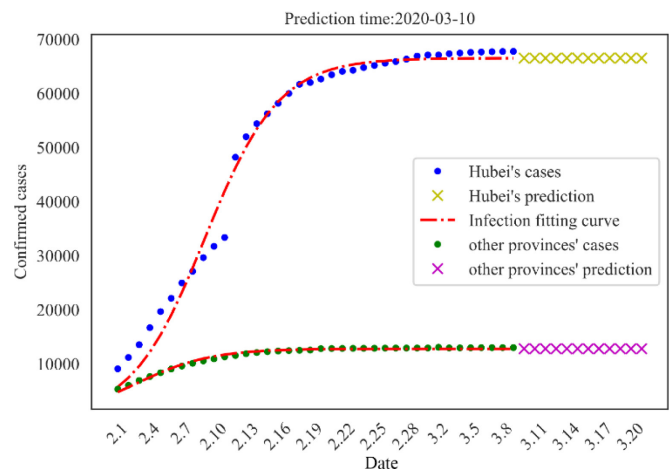
The number of confirmed cases in other provinces and cities has not increased basically since February 16, 2020. The critical point of Hubei Province emerges at March 1.

more accurate information. This process of infection in January is well represented by our model (Figure 4).

The virus has been spreading at a rate of r = 1.0 since February, and the infectious cases increased rapidly at this incubation period, especially on February 12, the infection cases jumped by 14,840 due to the change of diagnostic criteria, leading the total cases to over 60,000. Under such a grim situation, a lockdown of communities and streets was enforced in Hubei Province. The political and medical community has continuously been fighting to persist with containment efforts. Compared with Hubei province, the epidemic situation in other provinces and cities in China kept a relatively stable state in February, and the number of confirmed cases basically has not increased since February 16, 2020 (Figure 5), reaching even zero cases in early March. Our model shows that the number of confirmed cases in Hubei Province decreased continuously at the end of February with the relative error -9.31%, and the critical point emerging on March the 1st. As a new public health measure, communities and streets quarantine is able to halt human-to-human transmission dramatically.

Evaluation of basic regeneration number via SEIR model

The basic regeneration number (R$_0$) is a key indicator of transmissibility to explain the average number of secondary cases generated by previous infectious cases. R$_0$ > 1 means the cases will increase, while R$_0$<1 means the cases will disappear gradually [16]. According to the Chinese Center for Disease Control and Prevention, the incubation period of COVID-19 is 14 days, and in severe cases, the average duration from onset to hospitalization was 9.84 days in Wuhan [17]. The median incubation period for SARS viruses similar to the COVID-19 was 6.4 days [18], while Chan *et al.* [19] analyzed nine cases of COVID-19 who were confirmed early at the known incubation period, and found that the average incubation period was 5.1 days. Such a long incubation period presents a severe test for our health. Here we calculated the basic regeneration number, with generation time tracing by Lipsitch *et al* and Northeastern University in the United States, so as to provide support for containment measures and centralized medical treatment.

According to the analysis of SARS transmission data by Lipsitch *et al* [20], the average number of parameter T$_g$ was 8.4 days, and 10 days by predicted results from Northeastern University in the United States. Here, we carried out the analysis based on the values of T$_g$ = 8.4 and T$_g$ = 10 respectively, and

borrowed the parameter state from Zhou [14] since restricted information about the probability of a suspected case becoming a confirmed case is uncertain.

According to the literature [14], the first case of COVID-19 was reported in December 8, 2019, here we supposed it as the beginning time (t = 0). There were 830 confirmed cases and 1072 suspected cases reported in 31 provinces and cities by 24:00 on January 23, 2020. Some infected cases with symptoms have not been tested yet [21] according to the current condition of prevention and control to COVID-19. We suppose the probability of the suspected case transferring to infection is ρ. According to the real-time news of China Daily, there were 259 new cases confirmed and 1072 suspected cases that had been reported, so $\rho = 259/1072 \approx 0.24$. We updated the parameter ρ and include infection cases up to Y(46) = 830 + 1072 × 0.24 ≈ 1087.

Similarly, there were 79824 confirmed cases and 851 suspected cases reported in 31 provinces and cities by 24:00 on February 29, 2020. According to the real-time news of China Daily, there were 444 new cases confirmed and 1965 suspected cases reported on January 24, 2020, so $\rho = 444/1965 \approx 0.23$. We updated the parameter ρ and included infection cases up to Y(36) = 79824 + 851 × 0.23 ≈ 80020.

There were 80754 confirmed cases and 349 suspected cases reported in 31 provinces and cities by 24:00 on March 9, 2020. According to the real-time news of China Daily, there were 125 new cases confirmed and 715 suspected cases reported on March 1, 2020, so $\rho = 125/715 \approx 0.17$.

We deduced the basic regeneration number on the basis of the data of CCTV news via the above explained process (Table 2).

COVID-19 differs in terms of infectious period, transmissibility, and clinical severity from previous infectious diseases, such as SARS, MERS, and Smallpox. We compared the R$_0$ in different periods. The value of R$_0$ was under 3 before the lockdown of Wuhan city as the date was January 23, 2020, which seemed lower than Zika (4.0) [22], MERS (4.35) [23], and Smallpox (4.25) [24]. In the period of peak incidence, the range of R$_0$ was up to [4.8, 5.8], and the increased number of the confirmed cases was 78537 within 36 days. By interrupting all human-to-human transmission, the value of R$_0$ dropped dramatically since early March. The range was [2.23, 2.51] in the whole testing span, which shows that secondary cases generated by an index case was less than 3 under the condition of no considering the "super spreaders".

Conclusion and discussion

Based on the data of COVID-19 in China from January 13, 2020 to March 9, 2020, in this paper we adopted the spatial Moran index to deduce that COVID-19 has had a significant spatial correlation in China. The epidemic situation in Hubei province is more severe, and the current transmission still presents an increasing trend. Our data analysis suggested that the epidemic could be effectively controlled since March 1, 2020, which is consistent with the actual situation. The transmission capacity of COVID-19 in other provinces and cities in China was relatively weak, perhaps the stringent quarantine contributed to achieve this result. The range of basic regeneration number of COVID-19 decreased, indicating that the spread rate of the disease slowed down. We expect that the Chinese government will continue to strengthen epidemic prevention and control and make the basic regeneration number to below 1 as soon as possible. Through results of our statistical modeling analysis of COVID-19, which has become a public health emergency of international concern, we hope to provide valuable reference information for medical personnel and health policy makers who are preparing for or experiencing the outbreak of novel coronavirus disease.

Our paper pays more attention to analyze the confirmed cases of this outbreak from the perspective of spatial measurement and statistical modeling. Because nucleic acid testing is time-consuming and labor-intensive, and requires professional equipment and technical personnel, data bias was inevitable, especially in Wuhan, the city with the most severe epidemic. Here we calculated the death rate at different regional scales from February 1 to March 9, which is shown in Figure 6.

Comparing data of other cities and provinces to Wuhan city, some deviation existed in the statistical timeliness in Wuhan so it is inevitable that maybe bias occurred in our results. On February 12, 2020, the Government of Wuhan asked to close the passageways between residential communities and required residents

Figure 6. Mortality of several cities with severe infections and Hubei province (Wuhan, Huanggang, Xiaogan are the cities of Hubei province, where the majority of COVID-19 cases were reported in this province. We compared Hubei, except Wuhan, with the sum of other provinces about the rate of death, which is the lowest line. They all are below 4%, except Wuhan city).
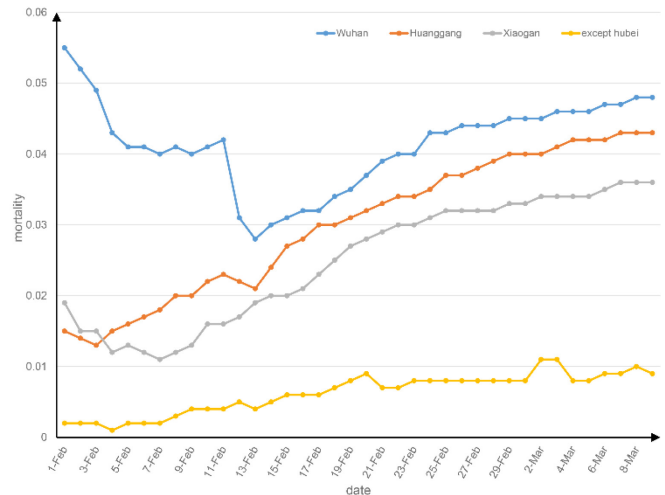
to stay at home. At the same time, the officers of residential community went from house by house to take residents' temperature, screen for fever or cough, to ensure that no suspected cases were missed. No significant difference between the statistical significance of Wuhan and other cities after that point, and the deviation of statistical information has been reduced (Figure 6). Our model shows that the epidemic situation of COVID-19 in Hubei Province can be effectively controlled based on the current prevention and control strategy. For those regions of out of Hubei province, the number of COVID-19 cases has decreased since February 16, and six provinces (Jilin, Hainan, Qinghai, Ningxia, Tibet and Gansu) have reported 0 cases, which is up to 29 provinces as of March 8. Strict quarantine has played a decisive role in preventing and controlling the virus.

Meanwhile, the prevention and control of COVID-19 is still a severe fight. On one hand, some questions are still uncertain and need to be answered, such as the

Table 2. Basic regeneration number.

Time node	T_g	ρ	λ	R_0
2019.12.8-2020.3.9	8.4	0.69	0.12	2.23
	10			2.51
2019.12.8-2020.1.23	8.4	0.24	0.16	2.19
	10			2.46
2020.1.24-2020.2.29	8.4	0.23	0.31	4.80
	10			5.80
2020.3.1-2020.3.9	8.4	0.17	0.12	2.15
	10			2.40

The dashed line is the separation between the whole timing span and the divided three stages.

identification of animal hosts, the period and route of transmission, the effective treatment and prevention methods (*i.e.* the effective test agents, drugs and vaccines), super spreaders, etc. On the other hand, workers return to their workplace except for Hubei province gradually. The rate of return to work is up to 90% in Tianjin (North China), 87.1% in Shanghai (Eastern China), 70% in Shanxi (North-East China) and 98.1% in Fujian (South China) until March the 6th. The mass flows of people greatly increase the risk of transmission of COVID-19 leading to detection and treatment of cases in the workplace difficult, the risk of another outbreak and rebound is still around China.

References

1. White House Coronavirus Task Force, Centers for Disease Control and Prevention (2020) Coronavirus Disease 2019 (COVID-19). Available: https://www.cdc.gov/coronavirus/2019-ncov/index.html. Accessed: 10 March 2020.
2. Michael Le Page, New Scientist (2020) COVID-19: The new coronavirus disease now officially has a name. Available: https://www.newscientist.com/article/2233218-covid-19-the-new-coronavirus-disease-now-off icially-has-a-name/#ixzz6DwhOEhpf. Accessed: 11 February 2020.
3. Miao Q, Cong XD, Wang B, Wang YG, Zhang ZD (2020) Understanding and thinking of novel coronavirus infection pneumonia in TCM. J Tradit Chin Med 61: 286-288.
4. Lu ZL, He RY, Wang WY, Fan T, Geng Q (2020) Clinical characteristics and immune function of patients with COVID-19. Med J Wuhan Univ 41: 1-5.
5. Tong S, Chen Z, Wu CY, Xu KY, Yang GH, Liao YD, Wang SH (2020) Surgical treatment of novel coronavirus (2019-nCoV) latent infection combined with esophageal foreign body perforation: a case report. Chin J Thoracic Cardiovasc Surg 27: 240-242.
6. Xiong B, Gou XY, Deng XY, Jiang JL, Peng J, Zhong HL, Zhang L (2020) The discuss of Novel coronavirus infectious pneumonia during the outpatient general hospital outpatient management. Chongqing Med J: 1-3.
7. Porcheddu R, Serra C, Kelvin D, Kelvin N, Rubino S (2020) Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. J Infect Dev Ctries 14: 125-128. doi: 10.3855/jidc.12600
8. Hussin AR, Siddappa NB The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun. In press.
9. Peng PWH, Ho PL, Hota SS Outbreak of a new coronavirus: what anaesthetists should know. Br J Anaesth. In press.
10. Han Q, Lin QQ, Jin SH, You LH Coronavirus 2019-nCoV: A brief perspective from the front line. J Infect. In press.
11. Ralph R, Lew J, Zeng T, Francis M, Xue B, Roux M, Toloue Ostadgavahi A, Rubino S, Dawe NJ, Al-Ahdal MN, Kelvin DJ, Richardson CD, Kindrachuk J, Falzarano D, Kelvin AA (2020) 2019-nCoV (Wuhan virus), a novel Coronavirus: human-to-human transmission, travel-related cases, and vaccine readiness. J Infect Dev Ctries 14: 3-17. doi: 10.3855/jidc.12425
12. Kelvin DJ, Rubino S (2020) Fear of the novel coronavirus. J Infect Dev Ctries 14: 1-2. doi: 10.3855/jidc.12496
13. Qi Y, Chen LH, Zhang L, Yang YY, Zhan SY, Fu CX (2020) Novel coronavirus infection in public: public cognition, attitude and behavior. J Trop Med 2: 1-12.
14. Zhou T, Liu QH, Yang ZM, Liao JY, Yang KX, Bai W, Lu X, Zhang W (2020) Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCov. J Evid Based Med 13: 3-7.
15. Zhao S, Lin QY, Ran JJ, Salihu S. Musa, Yang GP, Wang WM, Lou YJ, Gao DZ, Yang L, He DH, Maggie H. Wang (2020) 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 92: 214-217.
16. Tillett H (1992) Infectious diseases of humans: Dynamics and control. In Anderson RM, May RM, editors. Epidemiology and Infection. New York: Oxford University Press. 211-212.

17. China CDC (2020) The epidemiological characteristics of an outbreak of novel coronavirus diseases (COVID-19) in China. J Clin Epidemiol 41: 145-151.
18. Donnelly CA, Ghani AC, Leung GM, Hedley AJ, Fraser C, Riley S, Abu-Raddad JL, Ho LM, Thach TQ, Chau P, Chan KP, Lam TH, Tse LY, Tsang T, Liu SH, Kong JH, Lau EM, Ferguson NM, Anderson RM (2003) Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. Lancet 361: 1761-1766.
19. Jasper FC, Yuan SF, Kok KH, Kelvin KT, Chu H, Yang J, Xing FF, Liu JL, Cyril CY, Rosana WP, Hoi-Wah T, Simon KL, Kwok-Hung C, Vincent KP, Wan-Mui C, Jonathan DI, Jian-Piao C, Vincent CC, Honglin C, Christopher KH, Kwok-Yung Y (2020) A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 395: 514-523.
20. Lipsitch M, Cohen T, Cooper B (2003) Transmission dynamics and control of severe acute respiratory syndrome. Sci 300: 1966-1970.
21. Read JM, Bridgen JRE, Cummings DAT, Ho A, Jewell CP (2020) Novel coronavirus 2019-nCoV: Eearly estimation of epidemiological parameters and epidemic predictions. medRxiv 1: 20018549.
22. Lessler J, Chaisson LH, Kucirka LM, Bi Q, Granta K & Saljie H (2016) Assessing the global threat from Zika virus. Sci 353: 6300.
23. Majumder MS, Rivers C, Lofgren E, Fisman D (2014) Estimation of MERS-coronavirus reproductive number and case fatality rate for the spring 2014 Saudi Arabia outbreak: insights from publicly available data. PLoS Current Outbreaks. Available: http://currents.plos.org/outbreaks/index.html%3Fp=40801.html. Accessed: 18 December 2014.
24. Eichner M, Dietz K (2003) Transmission potential of smallpox: estimates based on detailed data from an outbreak. Am J Epidemiol 158: 110-117.

Corresponding author
Yongmei Ding, PhD
Associate Professor
Faculty of Statistics - Department of Mathematics and Statistics – Wuhan University of Science and Technology
No.22317, 2nd Building, Huangjiahu Campus,
Huangjiahu West Road, Hongshan District, Wuhan City, Hubei Province, P.R. of China, 430065
Tel: 1-806-258-9916
Email: dingyongmei@wust.edu.cn

Miao Liu, PhD
Faculty of Department of Pathology, Brigham and Women's Hospital, Harvard Medical School
77 Avenue Louis Pasteur, Boston MA, US, 02215
Tel: 1-857-350-7065
Mail: mliu0@bwh.harvard.edu

Conflict of interests: No conflict of interests is declared.