Prediction of Corona Virus Disease 2019 based on SIR model in China

jiangang mu (✉ qdbr2007@163.com)

Affiliated hospital of Qingdao university

Research

Keywords: Corona Virus Disease 2019, SIR model, Infective, removed, fit

Posted Date: August 21st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-58708/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Title:

**prediction of Corona Virus Disease 2019 based on SIR model
in China**

Authors:

Jiangang Mu MD.

Authors' addresses:

the affiliated hospital of Qingdao university, No.16 Jiangsu road , Qingdao 266003,
Shandong Province, China.

Corresponding author:

Jiangang Mu,

the affiliated hospital of Qingdao university,

No.16 Jiangsu road , Qingdao 266003, Shandong Province, China.

Tel: +86 13001670741

E-mail: qdbr2007@163.com

prediction of Corona Virus Disease 2019 based on SIR model in China

Abstract

Background: From the end of 2019 to the beginning of 2020, the Corona Virus Disease 2019 epidemic occurred in Wuhan, China, and quickly spread to the whole country. We used SIR epidemic model to predict the epidemic trend in China.

Materials and Methods: Respectively fitted the Corona Virus Disease 2019 epidemic trend equations in China, Hubei, and Wuhan, predicted future trends, based on the hypothesis of the infectious disease process by the SIR model and the official announcement data of the Chinese Health Commission .

Results: There will be no new cases in the non-Hubei area nationwide after March 8; there will be no new cases in Hubei that non Wuhan after March 12; Wuhan will there be no new cases after March 22 ;

Conclusions: The epidemic will end soon in China, under the prevention and control measures are not relaxed.

Keywords: Corona Virus Disease 2019, SIR model, Infective, removed, fit

Background

From late 2019 to early 2020, an epidemic of Corona Virus Disease 2019 (COVID-19) occurred in Wuhan, China [1,2], which was nearing the traditional Chinese New Year, and the mobility of people was relatively large, so the epidemic quickly spread to all provinces across the country. On March 13, 2020, China had accumulatively confirmed 81,029 patients, most of them in Wuhan, with 49,995 cases. As of March 13, 2020, a total of 66,990 cases had been confirmed in the world except China. They were distributed to five continents, and more than 100 countries [3,4]. In some countries, the outbreak is still on the rise stage. As the Chinese government has taken a series of measures in time to stop the spread of infectious diseases, including blocking Wuhan and Hubei, delaying the start of construction and schooling, the epidemic situation in China has been in a significant decline.

Since ancient times, infectious diseases have never stopped attacking human beings. It will threaten our health and life, trouble our lives, and have a serious impact on economic development. The fight between it with us have never stopped. In addition to studying how to eliminate it and how to prevent its spread, we also need to understand its development trend. People have established a variety of analysis models based on the characteristics of infectious diseases [5,6,7,8]. The transmission characteristics of COVID-19 are more in line with the SIR model. Therefore, this article uses the SIR model and China's COVID-19 epidemic data. Fitted the development curve equation of the epidemic, predicted its approximate end time in China.

Methods

The SIR model refers to dividing the population into three categories: Susceptible, Infective and Removed. The first letter of these three words combine called the SIR model. It was proposed by Kermack and McKendrick in 1927 [9]. Later, it was used by many scholars for the evaluation and analysis of infectious diseases [10,11,12,13,14,15]. The susceptible person refers to the susceptible population that has not yet been infected. Many infectious diseases such as COVID-19 and SARS can infect everyone. Susceptible here refers to all healthy people who are not infected. The infected person is the patient. He can infect healthy people and become patients. Removed refer to Removed refers to a patient who has been cured to gain immunity or died, they will no longer be infected and will withdraw from the infection system. This model is based on the following assumptions:

1. Suppose $s(t)$, $i(t)$, $r(t)$ respectively represent the proportion of Susceptible (healthy people), infected people (patients), and removed in the total population at time t in the area. The total population is a constant N , irrespective of natural birth rate and mortality, that is, $s(t) + i(t) + r(t) = 1$.
2. The average number of effective contacts per patient per day is a constant λ . λ is called the daily contact rate. When a patient is in effective contact with a healthy person, the healthy person becomes infected and becomes a patient. λ can also be called the daily infection rate.
3. μ is the removal rate, which is the proportion of removals that are cured and die from the patient daily.

The equation of the SIR model can be obtained:

$$\begin{cases} \frac{ds}{dt} = -\lambda si, & s(0) = s_0 > 0 \\ \frac{di}{dt} = \lambda si - \mu i, & i(0) = i_0 > 0 \\ \frac{dr}{dt} = \mu i, & r(0) = r_0 \geq 0 \end{cases} \quad (1)$$

In equation (1), s_0 is the proportion of healthy people at the initial time ($t = 0$), i_0 is the proportion of patients at the initial time, and r_0 is the proportion of people who are removed at the initial time. Regardless of the initial time s_0 , i_0 , the patient will eventually disappear, that is, $i = 0$. Because the final ending of infectious disease is either infecting everyone, or the infectious disease is cut off until it disappears. The model shows that the rate of change of the number of susceptible persons by time ($\frac{ds}{dt}$) is directly proportional to the product of the number of susceptible persons and the number of infected persons at that time. The rate of change of the number of removed by time is directly proportional to the number of infected persons and the rate of removed at that time.

Kermack WO and McKendrick AG in their article introduced the equation of change of the removed based on the model:

$$\frac{dz}{dt} = \frac{l^2}{2x_0 k^2} \sqrt{-q} \operatorname{sech}^2\left(\frac{\sqrt{-q}}{2} (t - \phi)\right) \quad (2)$$

$\frac{dz}{dt}$ in formula (2) is equal to $\frac{dr}{dt}$ in formula (1), which represents the change of

removal rate of patient, $k = \lambda$, $l = \mu$, $x_0 = s_0$,

$$\sqrt{-q} = \left\{ \left(\frac{k}{l} x_0 - 1 \right)^2 + 2x_0 y_0 \frac{k^2}{l^2} \right\}^{\frac{1}{2}} \quad (3)$$

$$\phi = \tanh^{-1} \frac{\frac{k}{l} x_0 - 1}{\sqrt{-q}} \quad (4)$$

Formula (2) was used to verify the plague in Mumbai, India in 1905 [9], and some scholars have used it to verify the SARS epidemic in Beijing in 2003.

COVID-19 has a latency period of several days [1,2,16], and humans are infective during the latency period. When a healthy person is infected, he is asymptomatic at first, and it is not clear that he has become an infected person. He will pass the disease to other Susceptible people that he contacts until he shows symptoms and seeks medical treatment. Isolation and treatment will not infect others. Therefore, it is difficult to determine the number of infected people in the population, but we can know the number of confirmed patients. The confirmed patients are equivalent to the removed, because once the patients are diagnosed, they will be isolated and treated, and there is almost no infectivity to the population. Although medical personnel have certain risks, we assume here that they are not infective.

Since the rate of change of removed is directly proportional to the number of infected persons, the daily rate of change of removed can represent the daily rate of change of infected persons. Therefore, according to the newly Confirmed diagnosis cases every day as change rate of removed, from January 20 to March 1, 2020, a total of 42 days of data, using Matlab software to fit the data curve according to formula (2), The fitting formulas of Wuhan, Hubei (excluding Wuhan) and the whole country (excluding Hubei) were obtained, and the end time of the epidemic was estimated

according to the formula. The above data comes from the official announcements of the Health Commission of China, Hubei and Wuhan, and a small part of the data comes from Baidu big data.

Results

Figure1 is the daily number of newly confirmed cases in China (excluding Hubei), and Figure 2 is a graph of epidemic trends in China except Hubei based on data fitting. Figure 3 is the daily number of newly confirmed cases in Hubei (excluding Wuhan), and Figure 4 is the epidemic trend figure of Hubei (excluding Wuhan) fitted according to the revised data of Figure3, Figure 5 is the daily number of newly confirmed cases in Wuhan. Figure 6 is a revised daily data figure of Figure5, and Figure 7 is a trend graph of epidemic trend in Wuhan fitted according to Figure6.

Among them, the data of Wuhan had been corrected because Wuhan City added 13,436 new cases on February 12, of which 12,364 were clinically diagnosed cases. Since diagnostic reagents were developed urgently after the outbreak, its production capacity may not be able to meet the rapid growth of patients. There was also insufficient testing staff for many patients, leading to many patients had not being diagnosed with the etiology, but they are extremely consistent with the COVID-19 in clinical indicators. In order to get them treated in time, the health department decided to adopt a clinical diagnosis method. These clinically diagnosed cases were not diagnosed at one time. They were centralized and treated before this day, so they lost their infectivity before that. They should have been confirmed diagnosis sometime before February 12th. (Moved out). Therefore, this article redistributes the clinically diagnosed cases on this day according to the weight of the number of new cases per day from January 20 to February 12. The Hubei data has also been revised, because a

total of 361 clinically diagnosed cases in Hubei were excluded from February 14 to February 19, resulting in a negative number of new cases in Hubei (Fig. 3). The number of cases was revised by weight from February 12 to February 17 (because nucleic acid tests usually take at least 2 days to produce results)

$$x=805.5072 \operatorname{sech}^2(0.1190t-1.7247) \quad (5)$$

$$x=1.1029e+03 \operatorname{sech}^2(0.1179t-1.9812) \quad (6)$$

$$x=2.8506e+03 \operatorname{sech}^2(0.1122t-2.3202) \quad (7)$$

Equation (5) is a national non-Hubei region equation fitted with Matlab software.

Equation (6) is an equation in Hubei without Wuhan, and equation (7) is an equation in Wuhan. According to the above equation, the number of new cases in non-Hubei areas nationwide can be calculated as 1.1079 on March 7, 0.8734 on March 8, 0.4278 on March 11, and 0.0502 on March 20. Increasing cases will tend to zero after March 8. The probability of new cases equaling zero on March 8 is 13%, 57% on March 11, and 95% on March 20. The non-Wuhan area of Hubei will return to zero after March 12, with a probability of 13% on March 12, 57% on March 15, and a 95% probability on March 24. New cases in Wuhan will return to zero after March 22, with a 14% probability that March 22 will be equal to zero, a 56% probability on March 25, and a 95% probability on April 4. The number of confirmed COVID-19 outbreaks in Wuhan is longer than in other parts of the country.

Discussion

The results in this article are calculated based on the SIR epidemic model, which may differ from the actual situation. No current model can fully predict the trend of infectious diseases, because there are many influencing factors, and these factors are dynamic. The daily infection rate (λ) and removal rate (μ) in the model are two important factors. When the infectious disease was first started, people were not aware of it, and there was no precaution. It was easy to be infected by the infected people, and the infection rate would be high. COVID-19 and other respiratory infectious diseases can be transmitted through droplets, contact, aerosol, etc. Way of transmission, so the initial infection rate is definitely high. When people know the existence of infectious diseases, precautionary measures will be taken naturally because of fear of the disease, and the government will also take necessary control measures to cut off the transmission route as much as possible, and the transmission rate will decrease. After the COVID-19 epidemic in Wuhan, the Chinese government adopted blockades in Wuhan and Hubei, and implemented grid-based management of residents from their communities of residence throughout the country. requiring people to wear masks, reduce the concentration of people, delayed construction and start school. As a result, the number of patients is declining. Asking everyone to put on a mask is not only a mask for healthy people, but also a mask for possible infected people, because infected people often don't know before the disease is diagnosed, and it may be contagious at that time. Medical conditions in a place can affect the rate of removal. Only when a patient is diagnosed can isolation treatment be removed from the infection system and no longer infect healthy people. Medical conditions are usually stable and the removal rate is only related to the number of patients. However, in the initial stage and outbreak of infectious diseases, there will be a shortage of medical resources. In the early COVID-19 epidemic in Wuhan, lack of masks,

protective clothing, wards, beds, medical staff, and nucleic acid testing reagents and personnel shortages will cause the epidemic expand. Later, these problems were resolved and the epidemic tended to ease.

In recent days, the epidemic has grown rapidly in many countries. As of April 3, the total number of confirmed cases worldwide has exceeded 1.01 million. A total of 115,242 cases have been confirmed in Italy, 112,065 in Spain, and 84,794 in Germany. the total number of confirmed cases in the United States has reached 245,573.

Now, there are still little new cases in China, but these cases are mainly caused by imported factors.

Conclusions

Based on the analysis of data and models for the COVID-19 epidemic in China, this article draws the following conclusions:

The number of new cases in China (excluding Hubei) will return to zero after March 8.

The number of new cases in Hubei (excluding Wuhan) will be reset to zero after March 12.

The number of new cases in Wuhan will return to zero after March 22.

The government can consider building reserve wards and beds in hospitals in response to public health emergencies.

Masks and protective clothing are important protective materials for respiratory infectious diseases.

The epidemic in China is about to end soon, but mathematical models have his limitations, and such predictions and data are based on the premise that current prevention and control measures are not relaxed. If there is some slack, the epidemic may be further expanded. It must be very different from the conclusion of this article.

Declarations:

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The datasets generated and/or analysed during the current study are available in the [National Health Commission of the People's

Republic of China] repository, [<http://www.nhc.gov.cn/>]

[<http://wjw.hubei.gov.cn/>][<http://wjw.wuhan.gov.cn/>]

Competing interests: The authors declare that they have no competing interests.

Funding: None .

Authors' contributions: The study conception and design, material preparation, data collection and analysis by M J , the first draft of the manuscript was written by M J. All authors read and approved the final manuscript.

Acknowledgements: Not applicable.

References

1. Huang C ¹, Wang Y ², Li X ³, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497-506.
2. Li Q¹, Guan X¹, Wu P¹, et al . Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med*. 2020 Jan 29. doi: 10.1056.
3. Burke RM, Midgley CM, Dratch A, et al. Active Monitoring of Persons Exposed to Patients with Confirmed COVID-19 - United States, January-February 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(9):245-246.
4. Bordi L ¹, Nicastri E ¹, Scorzolini L ¹, Di Caro A ¹, Capobianchi MR ¹, Castilletti C ¹, Lalle E ¹, On Behalf Of Inmi Covid-Study Group And Collaborating Centers ^{2, 1}. Differential diagnosis of illness in patients under investigation for the novel coronavirus (SARS-CoV-2), Italy, February 2020. *Euro Surveill*. 2020 Feb;25(8). doi: 10.2807/1560-7917.ES.2020.25.8.2000170.
5. Serfling RE. Methods for current statistical analysis of excess pneumonia-influenza deaths [J]. *Public Health Rep*, 1963, 78(6):494-506.
6. Simonsen L, Clarke MJ, Williamson GD, et al. The impact of influenza epidemics on mortality: introducing a severity index [J]. *Am J Public Health*, 1997, 87(12):1944-1950.
7. Choi K, Thacker SB. An evaluation of influenza mortality surveillance, 1962-1979. II. percentage of pneumonia and influenza deaths as an indicator of influenza activity[J]. *Am J Epidemiol*, 1981, 113(3): 227-235.

8. Yang L, Ma S, Chen YP, et al. Influenza associated mortality in the subtropics and tropics: results from three Asian cities [J]. *Vaccine*, 2011, 48(29): 8909-8914.
9. Kermack WO, McKendrick AG, A contribution to the mathematical theory of epidemics [J]. *proc Royal Soc (Series A)*, 1927;115:700-721.
10. Li GH¹, Zhang YX¹, Dynamic behaviors of a modified SIR model in epidemic diseases using nonlinear incidence and recovery rates. *PLoS One*. **2017**; 12(4):e0175789. doi: 10.1371/journal.pone.0175789.
11. M. Saeedian,¹ M. Khalighi,¹ N. Azimi-Tafreshi,² G. R. Jafari,^{1,3,4} and M. Ausloos^{5,6,7}, Memory effects on epidemic evolution: The susceptible-infected-recovered epidemic model. *Phys Rev E*. 2017;95(2-1):022409. doi: 10.1103.
12. Maltz A¹, Fabricius G², SIR model with local and global infective contacts: A deterministic approach and applications. *Theor Popul Biol*. 2016;112:70-79.
13. Cao Z¹, Liu X², Wen X³, Liu L⁴, Zu L⁵. A regime-switching SIR epidemic model with a ratio-dependent incidence rate and degenerate diffusion. *Sci Rep*. 2019;9(1):10696. doi: 10.1038/s41598-019-47131-6.
14. Kaminsky J¹, Keegan LT¹, Metcalf CJE^{1,2}, Lessler J¹. Perfect counterfactuals for epidemic simulations. *Philos Trans R Soc Lond B Biol Sci*. 2019 Jul 8;374(1776):20180279. doi: 10.1098/rstb.2018.0279.
15. Kleczkowski A¹, Hoyle A², McMenemy P². One model to rule them all? Modelling approaches across OneHealth for human, animal and plant epidemics.

Philos Trans R Soc Lond B Biol Sci. 2019 Jun 24;374(1775):20180255. doi:
10.1098/rstb.2018.0255.

16. Tian S ¹, Hu N ², Lou J ¹, et al . Characteristics of COVID-19 infection in Beijing.
J Infect. 2020 Feb 27. pii: S0163-4453(20)30101-8. doi: 10.1016/j.jinf.2020.02.018.

Fig1 is the daily number of newly confirmed cases in China (excluding Hubei).

Fig2 is a graph of epidemic trends in China except Hubei based on data of figure1 fitting.

Fig3 is the daily number of newly confirmed cases in Hubei (excluding Wuhan).

Fig4 is the epidemic trend graph of Hubei (excluding Wuhan) fitted according to the revised data of Fig3.

Fig5 is the daily number of newly confirmed cases in Wuhan.

Fig6 is a revised figure of the daily number of newly confirmed cases in Wuhan.

Fig7 is a trend graph of epidemic in Wuhan fitted according to Fig6.

Figure1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figures

Figure 1

Daily number of newly confirmed cases in China (excluding Hubei).

Figure 2

Graph of epidemic trends in China except Hubei based on data of figure1 fitting.

Figure 3

Daily number of newly confirmed cases in Hubei (excluding Wuhan).

Figure 4

Epidemic trend graph of Hubei (excluding Wuhan) fitted according to the revised data of Fig3.

Figure 5

Daily number of newly confirmed cases in Wuhan.

Figure 6

Revised figure of the daily number of newly confirmed cases in Wuhan.

Figure 7

A trend graph of epidemic in Wuhan fitted according to Fig6.