The effects of the clinical symptoms pneumonia-confirmation strategy of the COVID-19 epidemic in Wuhan, China

Yanjin Wang*, Pei Wang, Shudao Zhang, Hao Pan

Institute of Applied Physics and Computational Mathematics, District Haidian, 100094, China
*Corresponding author. Email: wang_yanjin@iapcm.ac.cn

Motivated by the quick control in Wuhan, China, and the rapid spread in other countries of COVID-19, we investigate the questions that what is the turning point in Wuhan by quantifying the variety of basic reproductive number after the lockdown city. The answer may help the world to control the COVID-19 epidemic. A modified SEIR model is used to study the COVID-19 epidemic in Wuhan city. Our model is calibrated by the hospitalized cases. The modeling result gives out that the means of basic reproductive numbers are 1.5517 (95% CI 1.1716-4.4283) for the period from Jan 25 to Feb 11, 2020, and 0.4738(95% CI 0.0997-0.8370) for the period from Feb 12 to Mar 10. The transmission rate fell after Feb 12, 2020 as a result of China’s COVID-19 strategy of keeping society distance and the medical support from all China, but principally because of the clinical symptoms to be used for the novel coronavirus pneumonia (NCP) confirmation in Wuhan since Feb 12, 2020. Clinical diagnosis can quicken up NCP-confirmation such that the COVID-19 patients can be isolated without delay. So the clinical symptoms pneumonia-confirmation is the turning point of the COVID-19 battle of Wuhan. The measure of clinical symptoms pneumonia-confirmation in Wuhan has delayed the growth and reduced size of the COVID-19 epidemic, decreased the peak number of the hospitalized cases by 96% in Wuhan. Our modeling also indicates that the earliest start date of COVID-19 in Wuhan may be Nov 2, 2019.

Abstract: COVID-19, Epidemic, Clinical Symptoms pneumonia-confirmation, Wuhan,

Introduce
In December 2019, a new coronavirus was firstly discovered in Wuhan, Province Hubei, China. The World health organization named this coronavirus as coronavirus disease 2019 (COVID-19) on Feb 11, 2020. In Jan 2020, COVID-19 epidemic began to break out in Wuhan (1). To control the COVID-19 epidemic, a serial strategy was taken. Firstly, a controversial measure was the lockdown of Wuhan and nearby cities in Province Hubei to prevent the further exportation of infected individuals to the rest of China. And then many health care teams were sent to support Wuhan from outside of Province Hubei, China. And anyone, who went outdoors, was asked to necessarily wear a mask. Especially, clinical symptoms begin to be used to confirm the novel coronavirus pneumonia (NCP) in the COVID-19 worst affected area of China from Feb 12, 2020, CT imaging results were used as the clinical diagnostic criteria for NCP, but strictly limited in Hubei Province (2). With the working of China’s strategy, until now the COVID-19 epidemic has been in control. And the number of daily new confirmed cases in Wuhan was dropped precipitously from thousands to single digital. China's COVID-19 strategy was successful. Chinese scientists also have made a massive effort in the battle of COVID-19 in Wuhan. But the COVID-19 epidemic has broken out in many other countries around the world. Now scientists are concerned about whether the China’s COVID-19 strategy can work elsewhere (2). To control the world-spread of the COVID-19 epidemic, it is necessary to understand what the key of the COVID-19 epidemic control is in Wuhan.

As we know, it is difficult to understand an epidemic in part because of a mass of uncertainty factors of individuals, institutions and government to a new and severe infecting disease. Over the last few decades, mathematical models of disease transmission have provided effective different intervention strategies for the infectious disease, and become a more and more powerful tool to analyze the epidemiological characteristics of infectious disease (4-7). The most common epidemic mathematical model is compartmental model. A compartmental approach always assumes that a susceptible individual first goes through a latent period (and be said to become exposed or in class E) after infection, would goes to the recovery. The resulting model is the SEIR model. The SEIR model has been used to study other two coronaviruses (SARS and MERS-Cov) since 2003(8-12). Also some researchers have studied COVID-19 by mathematical model (13-17). Except for one
study, which concerned the optimizing clinical diagnosis (18), as we know there is no result to study the effect of clinical symptoms NCP-confirmation.

Within computational science, uncertainty quantification refers to a family of solution family to characterize the variability of a given analysis and the spread in the predicted performance of a complex system. In the Bayesian framework, model parameters are considered as random variable and the uncertainty about them is updated using observed data. A likelihood function is necessary for Bayesian inference. But for complex models from medical and biological science the likelihood functions are computationally intractable. In this work approximate Bayesian computing (19-21) is used to calibrate the model parameters.

In this study, our proposal concentrates on the role of the COVID-19 epidemic control in Wuhan after the lockdown city, quantifying the importance of the clinic symptoms for NCP-confirmation. To this end, we calibrate the modified SEIR model by the public data. We quantify the variety of basic reproductive number before and after Feb 12, 2020. Furthermore, we made a long period modeling to show that the spreading of the epidemic.

RESULTS

In Wuhan, COVID-19 broke out since Jan 2020, and is under control in no more than two months. However it has been breaking out in other countries. We need to study the COVID-19 in Wuhan and other countries.

We first analyze the public data. Figure 1A indicates that there is jumping on Feb 12. Thus we separate the period into two stages: the 1st stage from Jan 25 to Feb 11, the 2nd stage from Feb 12 to Mar 10. And the data from Mar 11 to 15 is used to validate our model. And then we calibrate our proposed mathematical model by public data. Figure 2A and B summarize that the fitness of the model and the public data. Also, the basic reproductive number was studied in Figure 2C and D. Since the basic reproductive number is 0.4738 in the 2nd stage, it is indicated that the COVID-19 epidemic has been under control in Wuhan. The measure of clinical symptoms pneumonia-confirmation in Wuhan has delayed the growth and reduced size of the COVID-19 epidemic, decreased the peak number of the hospitalized cases by 96% in Wuhan. Figure 2B tells us that the hospitalized cases will be clear as early as the end of April, 2020. At last we find that the earliest transmission of the COVID-19 epidemic in Wuhan began on Nov 2, 2019.

MATHEMATICAL METHOD

Data

The observed data is collected from the official public data of Wuhan Municipal Health Commission (22) and the big data of BAIDU (23). To obtain reliable observed data, we use the data of the numbers of daily confirmed, recovered and death cases from February 25 to March 15 after the lockdown of Wuhan city. There is jumping on February 12 in Fig. 1A. As we know that on that day the clinical symptoms were used to confirm NCP besides the laboratory confirmation in Wuhan (2, 24). So we would separate the considered period into two stages: January 25 to February 11 (denoted as the 1st stage) and February 12 to March 15 (denoted 2nd stage) to see the epidemic trend for-and-aft in Feb 12.

Mathematical Model

Many researchers have made significant progress on the SEIR model, where S, E, I, R denote mutually exclusive classes containing susceptible, exposed (latent), infectious, and recovered individuals. As the different epidemics have different transmission characteristics, the epidemic mathematical model cannot straightforwardly use. We class the population into susceptible, exposed, infectious, hospitalized and removed (Fig. 1B). Our model also considers that the variety of the transmission affects the daily rate of spread from the susceptible class to exposed class. Based on the SEIR model we propose a modified SEIR model for the transmission of people between five states: susceptible (S), exposed (E), infected (I), hospitalized (H) and removed (R), where the infected class includes the infections with symptoms and infections but no symptoms, and removed class includes the recovery and death without considering the nature birth and death, our COVID-19 epidemiological model, is given by the following nonlinear system of ordinary differential equations:
which is denoted as SEIHR with the initial values 
\((S_0, E_0, I_0, H_0, R_0)\), satisfying the following
conservation law of the population
\[ N = S(t) + E(t) + H(t) + R(t), \]

Here \(N\) means the total population without considering
the number of natural birth and death, and the meaning of the
parameters sees Table 1.

| Parameter | Definition |
|-----------|------------|
| \(\beta(1-I/N)\) | Transmission rate per day |
| \(\lambda\) | Rate of progression to infectious state per day |
| \(\alpha\) | Rate of progression from infections to hospitalized cases per day |
| \(\gamma\) | Rate of from infections to removed state per day |
| \(\mu\) | Rate of from the hospitalized cases to removed cases per day |

Figure 1: The used public data and model structure. (A) Daily of the
hospitalized cases in Wuhan. (B) Flow-SEIHR model diagram.
Susceptible individuals (S) become exposed persons (E) before
infections (I). After infections they then progress to two classes:
hospitalized cases with confirmation COVID-19 patients and
removed class without non-confirmation including recovery and
death. Hospitalized case finally enters the removed class. (C) The
flow chart to study an epidemic model

From the second generator approach \(^{26,27}\), we obtained the
control basic reproductive number as following
\[
R_0 = \frac{\beta}{\alpha + \gamma} \frac{S_0}{N} \approx \frac{\beta}{\alpha + \gamma}
\]

By the posterior PDF of the parameters, we may easily
deduce the PDF of \(R_0\) with means and variance.

To calibrate the parameters, there are usually two ways:
Maximum likelihood and Bayesian analysis. For both
approaches the likelihood function is needed. But
appropriate likelihood function is usually difficult.
Approximate Bayesian Computing (ABC) \(^{18-20}\) is a set of
methods that attempts to use Bayesian idea without the
likelihood function. In this study we use ABC to estimate
the parameters. Analyzing the data we see that the
parameter, \(\mu\), was no more than 0.1. We assume that \(\mu\)
submitted to the uniform PDF on \([0, 0.1]\). Except \(\mu\), all the
other parameters are assumed to submit the uniform PDF
on \([0, 1]\). Firstly, the parameters are sampled in the
uncertainty intervals 50000 times by Latin hypercube
design \(^{28-30}\). And then we define the following criteria for
accepted samples
\[ d = \sqrt{\frac{\sum_{t=1}^{Days} \left( \ln(H(t)) - \ln(H_{obs}(t)) \right)^2}{\text{Days}}}, \]
\[ \leq \varepsilon \sqrt{\frac{\sum_{t=1}^{Days} \left( \ln(H_{obs}(t)) \right)^2}{\text{Days}}} \]

where \( \varepsilon \) is a small positive number.

After we obtain the accepted samples, the corresponding statistical values may be deduced. For every sample of parameters we use MATLAB to solve the ordinary differential system of the SEIHR model. The flow chart to study an epidemic model is shown in Fig. 1C.

**MODELING THE COVID-19 EPIDEMIC IN WUHAN**

This part focuses on the spread of COVID-19 epidemic in Wuhan after the lockdown of this city. We mainly concern the number of hospitalized infections, which means the confirmed infections except for recovery and death.

By approximate Bayesian computing \(^{18-20}\) we calibrate our model \(^{25}\). We set the modeling result of Feb. 12 with the means of the parameters for the 1\(^{st}\) stage as the initial values of the model of the 2\(^{nd}\) stage, except that the number of hospitalized infections is adjusted to the public data on Feb 12 and the number of infections is also adjusted. Our model fits to the case-incidence data qualitatively well (Fig. 2A for the 1\(^{st}\) stage and Fig. 2B for the 2\(^{nd}\) stage in Wuhan). For these two stages, 95% CIs of the hospitalized cases (the grey region in Fig. 2A and B) cover the public data, and the modeling results with the corresponding means of the parameters are coherent with the observed data. By the sampling modeling we then estimate that the means of \( R_0 \) are 1.5517 with %95 CI [1.1716, 4.4283] for the 1\(^{st}\) stage (Fig. 2C), and 0.4738 with 95% CI [0.0997, 0.8370] for the 2\(^{nd}\) stage in Wuhan (Fig. 2D). These results indicate \( R_0 \) is below the self-sustaining threshold of 1 since Feb 12, 2020, in Wuhan, thus, the Wuhan epidemic has been under control, in the sense that infection rates are declined. We notice that Feb 12 exactly is the date when the clinical diagnosis began to confirm NCP in Wuhan. Fig. 2B illuminates that the hospitalized cases in Wuhan may be cleared as early as the end of April, 2020.

By the public data in Wuhan, we see that the peak number of hospitalized cases is 37,755 on Feb 18, 2020. But if we suppose that there is no clinical symptom confirming NCP, taking a long time simulation with the means of the parameters for the 1\(^{st}\) stage of Wuhan we estimate that the peak of hospitalized cases is the 74\(^{th}\) day since Jan 25, 2020, that is to say, it is on April 8. The peak number of
hospitalized cases was 979966. The peak number of the hospitalized cases was dropped by 96% (Fig. 3A). Furthermore, under the assumption that at the start time of COVID-19 with two latent, one infectious and zero hospitalized cases, we modeled the COVID-19 epidemic for a long time transmission. The result shows that the peak time of hospitalized cases is the 145th day, the peak number was 976472 (Fig. 3B). We find that the above two peak numbers are almost the same. We also find that, the date, which was mostly closed to the public data on Jan 25, 2020, is the 56th day after the start time of COVID-19. The modeling result of hospitalized infectious is 560 (the public data was 533) at the red line of Fig. 3B. Furthermore, if we suppose that in the start date the initial exposed class has only one case, and the infectious, hospitalized and removed classes are zero cases, then the peaking data of the hospitalized cases is the 157th day and the mostly close date to the public data on Jan 25, 2020 is the 66th day with the hospitalized number, 509.

**DISUSSION**

As soon as the COVID-19 epidemic can transmit human to human, the infection number would rise exponentially without control measures. Fortunately, China’s government has taken a highly crucial strategy to control COVID-19 in Wuhan. In this study we have simulated the transmission dynamics of COVID-19 for the different stages in Wuhan. And we show that China’s clinical diagnosis strategy is all-important and effective for the epidemic control. The experience maybe helps the world to control the COVID-19 epidemic.

By the long period modeling of the 1st stage of Wuhan city we estimate the peaking date of hospitalized cases is April 8 and the peak number of hospitalized is more than 900 thousand. If it happens, so many infective patients must overwhelm the public health service. The public data shows that the peaking number of the public data of hospitalized cases is 37755 on February 18, 2020. Comparing the public data and the above modeling peaking number of the hospitalized cases, we see that the number dropped by 96% and the period was shorten two-part. China’s clinical symptom NCP-confirmation strategy shortens the period of the COVID-epidemic. In another words, clinical symptom NCP-confirmation strategy make to find the infections as soon as possible, China’s other strategy, building Leishenshan, Huoshenshan and Fangcang hospitals and the health care teams from other province of China, makes the confirmed infections to be hospitalized as soon as possible. The result shows that on February 12 China’s strategy is the key of the COVID-19 battle in Wuhan. From the mathematical point of view, it is also seen that, if suddenly a large number of infectious cases turned into the hospitalized cases, then the spread of the COVID-19 epidemic will change.

**Table 2. The estimation of start data of COVID-19 epidemic in Wuhan**

| Start Date    | Evidencty      | Source       |
|---------------|----------------|--------------|
| Dec 8, 2019   | Clinical cases | *N Engl Med J* (1) |
| Dec 1, 2019   | Clinical cases | *Lancet* (31) |
| Nov, 2019     | Inferring from the above result | *Science* (32) |
| Nov 15, 2019  to Nov 30, | Modeling with $(E_0,I_0,H_0,R_0) = (2,1,0,0)$ | *This study* |
For a long period of modeling with the parameters of the 1st stage, we use different manners to estimate the start time. One estimated start date is Nov 15, 2019, and another estimated start date is Nov 30, 2019. Our estimating start date is in good agreement with the pioneering results in Table 2. It is amazing that the period between the two dates was almost an isolation cycle. The period between two dates was also consistent with the preliminary estimate of the serial interval (mean 7.5 days and 95% CI 5.3 to 19). Making furthermore assumption, we see that the estimated start date of the COVID-19 epidemic in Wuhan is from Nov 2 to Nov 20, 2019.

In this work, to understand the transmission dynamics of the COVID-19 epidemic in Wuhan and other countries, we quantify the effectiveness of the COVID-19 epidemic control in Wuhan. By the public data we considered the period as two stages in Wuhan. The model was calibrated by the public data. The results show that Feb 12 is the sticking point of the Wuhan epidemic control. As different measures are always coupling together, this study drew inferences from statistical and mathematical model. The mathematical models are complex, the data are poor, and some big questions remain. Misdiagnosis are almost inevitable by the clinical symptom NCP-confirmation, could affect measurements of the rate of epidemic spread. Mathematical model cannot consider every aspect, so there are still some uncertainties in the prediction of the epidemic transmission. Figure 2B shows that the hospitalized cases in Wuhan may be clear as early as the end of April, 2020. But with the decreasing of the hospital cases more and more medical resource focuses the hospitalized cases such that the clear date in Wuhan may be earlier than our modeling prediction.

Reference

1. Li Q, Guan XH, Wu P, et al. Early transmission dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med. DOI: 10.1056/NEJMoa2001316.
2. National Health Commission of The People’s republic of China. Diagnosis and treatment guideline for new coronavirus pneumonia(revised version fifth). Feb 8 2020.
3. Kupferschmidt K, Cohen J, Can China’s COVID-19 strategy work elsewhere. Science 2020, 367(6482):1061-1062.
4. Anderson RM, May RM. Infectious disease of humans. Oxford: Oxford University Press; 1991.
5. Heesterbeek H, Anderson RM, Anderson V, et al. Modeling infectious disease dynamics in the complex landscape of global health. Science 2015; 347: aaa4339.
6. Ross R. The prevention of malaria. London: John Muarry; 1911.
7. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics, IV: analysis of experimental epidemics of the virus disease mouse ectromelia. J Hyg(Lond) 1939;37:172-87.
8. Ghowell G, Fenimore PW, Castillo-Garsow MA, Castillo-Chavez C. SARS outbreaks n Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism. J. Theo. Bio. 2003, 224:1-8.
9. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of Public Health Interventions. Science 2003, 300:1961-1966.
10. Dye C, Gay N. Modeling the SARS epidemic. Sience 2003, 300: 1884-1885.
11. Lipsitch M, Cohen T, Cooper B, robins JM, Ma S, James L. Transmission dynamics and control of severe acute respiratory syndrome. Sience 2003, 300: 1966-1970.
12. Chowell G, Blumberg S, Sirmonsent L, Miller MA, Viboud C. Synthesizing data and models for the spread of MERS-Cov, 2013: key role of index cases and hospital transmission. Epidemic, 2014, 9:40-51.
13. 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 modeling study. Lancet 2020, DOI: 10.1016/S0140-6736(20)30260-9.
14. Yang Z, Zeng Z, Wang K, et al. Modified SEIR and prediction af the epidemics tren of COVID-19 in China under public health interventions. J. Theoac. Dis. 2020, 12(3): 165-174.
15. Chinazzi M, Dvis JT, Ajelli Marco. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreaks. Science, 2020. DOI: 10.1126/science.ab9757.
16. Yuan Z, Yuan C. Simple model to assess Wuhan lock-down effect and region efforts during COVID-19 epidemic in China mainland. MedRxiv. 2020. DOI: 10.1101/2020.02.29.20029561.
17. Zhang B, Zhou H, Zhou F. Study on SARS-COV-2 transmission and the effects of control measures in China. medRxiv, 2020. DOI: 10.1101/2020.02.16.20023770.
18. Ai J, Zhang H, Xu T, Wu J, Zhu M, Yu YQ, et al. Optimizing diagnostic strategy for novel coronavirus pneumonia, a
multi-center study in eastern China. medRxiv 2020, DOI: 10.1101/2020.02.13.20022673.

19 Beaumont MA, Cornuet JM, Marin JM, Robert CP. Adaptive approximate Bayesian computation. Biometrika 2009, 96, 983-990.

20 Fearnhead P, Prangle D. Constructing summary statistic for approximate Bayesian computation: semi-automatic approximate Bayesian computation. J.R. Stat. Soc. Ser. B Stat. Methodol. 2012, 74:419-474.

21 BN Vo, CC Drovandi, AN Pettitt, MJ Simpson. Quantifying uncertainty in parameter estimates for stochastic models of collective cell spreading using approximate Bayesian computation. Mathematical Biosciences 263(2015):133-142.

22 Wuhan Municipal Health Commission, Hubei Province, http://wjw.wuhan.gov.cn.

23 Baidu. Real time data of novel pneumonia epidemic situation. http://voice.baidu.com/act/newpneumonia.

24 National Health Commission of The People’s republic of China. Up to 24:00 February 12th novel coronavirus pneumonia epidemic situation. Feb 13, 2020.

25 See supplementary materials.

26 Diekmann O, Heesterbeek JAP. Mathematical Epidemiology of infectious diseases: model building, analysis and interpretation. Wiley, New York 2000.

27 Towers S, Geisse KV, Zheng Y. Assessing potential repercussions using a seasonally forced SIR model. Journal of Theoretical Biology 2011, 289:259-268.

28 Tang, B. Orthogonal array-based Latin hypercubes.” Journal of the American Statistical Association, 1993, 88: 1392[1397.

29 Owen, AB. Controlling correlations in Latin hypercube samples. Journal of the American Statistical Association, 1994, 89: 1517-1522.

30 Linkletter C, Bingham D, Sitter R, Ye K, Hengartner N, Higdon D. Comparing Designs of Experiments for Screening and Prediction in Computer Experiments. Technical Report LA-UR-03-8524, Los Alamos National Laboratory, 2003.

31 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.

32 Cohn J, Wuhan seafood market may not be source of novel virus spreading globally. Science 2020, DOI: 10.1126/science.abb0611.

Funding: NSF(Grant No. 11671413) of China and the Science Challenge Project (Grant No. TZZ2018001).

Author contributions
Yanjin Wang, Pei Wang & Shudao Zhang conceived of this work and the analyzed the result; Yanjin Wang & Hao Pan carried out the collection of public data and the modeling.

Competing interests
The authors declare no competing interests.

Data
The data of the COVID-19 epidemic in Wuhan, collected in the current study, is publicly available, which can be found on official website of Wuhan Municipal Health Commission (Ref. 22) and the big data of Baidu (Ref. 23).