Real-time estimation of the reproduction number of the novel coronavirus disease (COVID-19) in China in 2020 based on incidence data

Kai Wang, Shi Zhao, Huling Li, Yateng Song, Lei Wang, Maggie H. Wang, Zhihang Peng, Hui Li, Daihai He

Department of Medical Engineering and Technology, Xinjiang Medical University, Urumqi 830011, China; JCS School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong 999077, China; Shenzhen Research Institute of Chinese University of Hong Kong, Shenzhen 518060, China; College of Public Health, Xinjiang Medical University, Urumqi 830011, China; Department of Epidemiology and Biostatistics, Department of Biostatistics, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China; Central Laboratory of Xinjiang Medical University, Urumqi 830011, China; Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong 999077, China

Contributions: (I) Conception and design: K Wang, H Li, D He; (II) Administrative support: Z Peng, MH Wang, D He; (III) Provision of study materials or patients: Z Peng, L Wang; (IV) Collection and assembly of data: S Zhao, Y Song, H Li; (V) Data analysis and interpretation: K Wang, S Zhao, MH Wang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Hui Li, PhD. Central Laboratory of Xinjiang Medical University, 393 Xinyi Road, Urumqi 830011, China. Email: huihui922@126.com; Daihai He, PhD. Department of Applied Mathematics, Hong Kong Polytechnic University, 11 Yuk Choi Road, Hung Hom, Kowloon, Hong Kong, China. Email: daihai.he@polyu.edu.hk.

Background: Since the first appearance in Wuhan, China in December 2019, the novel coronavirus disease (COVID-19) has posed serious threats to the public health in many Chinese places and overseas. It is essential to quantify the transmissibility on real-time basis for designing public health responses.

Methods: We estimated the time-varying reproduction numbers in China, Hubei province and Wuhan city by using the renewable equation determined by the serial interval (SI) of COVID-19. We compare the average reproduction numbers in different periods of time to explore the effectiveness of the public health control measures against the COVID-19 epidemic.

Results: We estimated the reproduction numbers at 2.61 (95% CI: 2.47–2.75), 2.76 (95% CI: 2.54–2.95) and 2.71 (95% CI: 2.43–3.01) for China, Hubei province and Wuhan respectively. We found that the reproduction number largely dropped after the city lockdown. As of February 16, the three reproduction numbers further reduced to 0.98, 1.14 and 1.41 respectively.

Conclusion: The control of COVID-19 epidemic was effective in substantially reducing the disease transmissibility in terms of the reproduction number in China reduced to 0.98 as of February 16. At the same time, the reproduction number in Wuhan was probably still larger than 1, and thus the enhancement in the public health control was recommended to maintain.

Keywords: COVID-19; outbreak; modelling; reproduction number; serial interval

Submitted Feb 26, 2020. Accepted for publication Apr 21, 2020.
do: 10.21037/atm-20-1944
View this article at: http://dx.doi.org/10.21037/atm-20-1944

© Annals of Translational Medicine. All rights reserved. Ann Transl Med 2020 | http://dx.doi.org/10.21037/atm-20-1944
Introduction

The novel coronavirus disease (COVID-19) cases have appeared in Wuhan, China since December 2019, while the causes of the first 4 cases are related to the South China Seafood Wholesale Market in Wuhan. Through the large-scale travelling activities, the COVID-19 has soon spread to other Chinese cities and other countries (Japan, Singapore, Thailand, etc.) (1,2). The COVID-19 could result in main symptoms including malaise, fever, dry cough, nasal congestion, diarrhea, and dyspnea, etc. (3). By examining the genome sequence of the virus, the pathogen of COVID-19 was found phylogenetically similar to the severe acute respiratory syndrome coronavirus (SARS-CoV), which caused outbreaks in China in 2003. Thus, the pathogen of COVID-19 was named as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the World Health Organization (WHO).

In order to reduce the risk of infection, the local government in Wuhan issued a notice that all within-city and cross-border public transportations were suspended as of January 23 (effect on 10 AM, GMT+8). Since then, the other 12 cities in Hubei province also implemented similar public traffic control measures. On January 30, the WHO declared that the COVID-19 outbreak constitutes a Public Health Emergency of International Concern (PHEIC) (4). Recently, studies have shown that the disease could be transmitted from person to person through droplets and contact (5). As of February 16, 2020, China has accumulated 70,548 confirmed cases, 1,770 dead cases, and 10,844 cured and discharged cases, of which Hubei province has the most severe epidemic situation, with a total of 58,182 confirmed cases (41,152 in Wuhan), accounting for 82.47% of the total death cases. The provinces with more than 1,000 confirmed cases in China are Hubei, Guangdong, Zhejiang, Henan, and Hunan (Figure 1).

The quantification of transmissibility—during epidemic is essential to design and adjust public health responses. Transmissibility can be measured by the reproduction number, the average number of secondary cases caused by an infected individual (6-9). Knowledge of the reproductive number is crucial for understanding the dynamics of any infectious disease, and these should be reevaluated as the pandemic progresses in space and time. Recently, the estimations of reproduction number of COVID-19 have been discussed in (1,2,10-14). The existence of these different findings is due to analyzing the problem from different perspectives. These authors selected datasets at different intervals, and the methods and parameters are also different. These different modeling methods have a bit varied estimates of reproduction numbers and also reflect the spread of coronavirus outbreaks and the dynamics of case reports differently. These findings provide important parameters for further analysis, including the impact of control measures and predicting the spread of future infections. Wu et al. (2) carried out a modeling study and estimated that the basic reproductive number for COVID-19 at 2.68 (95% CI: 2.47–2.86) from 1 December 2019 to 25 January 2020. Based on the cumulative number of reported cases data from 10 January 2020 to 22 January 2020 of Wuhan and Hubei province, Tang et al. (10) have estimated the basic reproduction numbers of Wuhan using likelihood-based and model-based approaches. The mean control reproduction number was estimated to be as high as 6.47 (95% CI: 5.71–7.23). Zhao et al. (11) adopted a data-driven modelling analysis of the Early Outbreak to estimate the number of unreported cases and the basic reproduction number of COVID-19 in Wuhan. They modeled the epidemic curve of COVID-19 cases, in mainland China from 1 December 2019 to 24 January 2020 through the exponential growth. After taking into account of the effect of under-reporting, the R0 was estimated at 2.56 (95% CI: 2.49–2.63). Meanwhile, Zhao et al. (12) modeled the epidemic curve of COVID-19 cases time series, in mainland China from January 10 to January 24, 2020, through the exponential growth. They estimated R0 by using the serial intervals (SI) of two other well-known coronavirus diseases, MERS and SARS, as approximations for the truly unknown SI. Their results showed that the mean R0 ranges from 2.24 (95% CI: 1.96–2.55) to 3.58 (95% CI: 2.89–4.39) with 2-fold to 8-fold increase in the reporting rate. Li et al. (1) collected information on demographic characteristics, exposure history, and illness timelines of laboratory-confirmed cases of NCIP (with 425 confirmed patients), that had been reported from January 1, 2020, to January 22, 2020. Using the data, they analyzed the early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. In the early period of exponential growth, they estimated the basic reproductive number was 2.2 (95% CI: 1.4 to 3.9). The dynamics transmission model by Shen et al. (13) computed a basic reproduction number of 4.71 (95% CI: 4.50–4.92), which decreased to 2.08 (95% CI: 1.99–2.18) on January 22nd, 2020. Finally, Riou et al. (14) computed a reproduction number of 2.2 (90% CI: 1.4–3.8)
Cumulative cases of COVID-19

![Cumulative cases of COVID-19](image)

**Figure 1** The geographic information map about cumulative number of confirmed COVID-19 cases as of February 16, 2020 in China.

by using a stochastic model.

With the implementation of COVID-19 control measures, the disease transmissibility was likely to change as well. Based on the surveillance data of COVID-19, we calculated the time-varying reproduction number, evaluated its changing dynamics from January 16 to February 16. Meanwhile, the effects of the COVID-19 control measures were evaluated by comparing the reproduction numbers in different periods of time.

**Methods**

We collected the data of the confirmed COVID-19 cases from the National Health Commission of the People's Republic of China from January 10 to February 16, 2020 (15). All cases were positively diagnosed based on the laboratory nucleic acid test, which is not excluded from the clinical diagnosis case. The time series plot of the daily number of COVID-19 cases is shown in **Figure 2**.

In this study, we divided the timeline of the outbreak into three stages:

- Stage I was from January 16 to January 22 before the ‘city lockdown’ was implemented in Wuhan;
- Stage II was from January 23 to January 30 before the WHO announced PHEIC; and
- Stage III was the time period after the PHEIC announced by WHO from January 31st to February 16.

We estimated the time-varying reproduction number, also known as the instantaneous reproduction number, by using the renewable equation proposed in Wallinga and Teunis (16). The relative likelihood $p_{ij}$ of case $i$ infected by case $j$, given their difference in time of symptom onset at $(t_i - t_j)$, can be expressed in terms of probability distribution for SI, denote by $\omega(t)$. Thus, the $p_{ij}$ is the likelihood that case $i$ has been infected by case $j$, normalized by the likelihood that case $i$ has been infected by any other case $k$:

$$p_{ij} = \frac{\omega(t_i - t_j)}{\sum_{ik} \omega(t_i - t_k)}$$  

[1]

The effective reproduction number for case $j$ is the sum of all cases $i$, weighted by the relative likelihood that case $i$ has been infected by case $j$:

$$R_j = \sum_i p_{ij}$$  

[2]

Additional detail on the derivation of these equations is provided in (17-19). The analysis was carried out by using R (version 3.6.0).

The SI is defined as the time between infections in a consecutive chain of transmission (20). The SI determines, along with the reproduction number, the growth rate of an epidemic curve. In You et al.'s study (21), the SI of COVID-19 was estimate at 4.4 days with a SD at 3.17 days based on 71 records. Thus, in this study, the SI is governed by a Gamma distribution with a mean of 4.4 days and a standard deviation of 3.17 days. Additionally, we examined
alternative scenarios in Tables S1 and S2 with SI at 7.5 days with a SD at 3.4 days, which was based on 5 pairs in Li et al.’s study (1).

**Results**

We estimated the time-varying reproduction numbers \( R(t) \) for the China as a whole, Hubei province and Wuhan respectively. The estimates of \( R(t) \) were shown in Figure 2. The estimated transmissibility appeared higher in the early stage of the epidemic. Then, in stage II, the \( R(t) \) gradually decreased. By February 15, the reproduction number dropped to 0.72 (95% CI: 0.62–0.82), 0.77 (95% CI: 0.67–0.88) and 0.90 (95% CI: 0.77–1.03) for China, Hubei and Wuhan respectively.

The average reproduction numbers of three stages in China, Hubei and Wuhan are shown in Table 1. In stage I, the reproduction number in China, Hubei province and Wuhan was 4.95 (95% CI: 4.26–5.67), 4.29 (95% CI: 3.66–4.93) and 3.88 (95% CI: 3.30–4.49) respectively. In stage II, the three reproduction numbers dropped to 2.61 (95% CI: 2.47–2.75), 2.76 (95% CI: 2.54–2.95) and 2.71 (95% CI: 2.43–3.01) respectively.
Table 2 The effect of under-report rate in the early outbreak on the average daily reproduction number with different stages in China, Hubei and Wuhan

<table>
<thead>
<tr>
<th>Under-report rate</th>
<th>Period</th>
<th>China</th>
<th></th>
<th>Hubei</th>
<th></th>
<th>Wuhan</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>95% CI</td>
<td>R</td>
<td>95% CI</td>
<td>R</td>
</tr>
<tr>
<td>20%</td>
<td>Stage I</td>
<td>3.40</td>
<td>(3.09, 3.70)</td>
<td>2.83</td>
<td>(2.52, 3.14)</td>
<td>2.34</td>
<td>(2.06, 2.63)</td>
</tr>
<tr>
<td></td>
<td>Stage II</td>
<td>2.10</td>
<td>(2.01, 2.19)</td>
<td>2.33</td>
<td>(2.20, 2.47)</td>
<td>2.45</td>
<td>(2.21, 2.69)</td>
</tr>
<tr>
<td></td>
<td>Stage III</td>
<td>1.02</td>
<td>(0.98, 1.06)</td>
<td>1.11</td>
<td>(1.07, 1.16)</td>
<td>1.31</td>
<td>(1.24, 1.38)</td>
</tr>
<tr>
<td>30%</td>
<td>Stage I</td>
<td>3.35</td>
<td>(3.06, 3.63)</td>
<td>2.77</td>
<td>(2.48, 3.06)</td>
<td>2.24</td>
<td>(1.98, 2.50)</td>
</tr>
<tr>
<td></td>
<td>Stage II</td>
<td>2.02</td>
<td>(1.94, 2.10)</td>
<td>2.24</td>
<td>(2.11, 2.36)</td>
<td>2.32</td>
<td>(2.11, 2.54)</td>
</tr>
<tr>
<td></td>
<td>Stage III</td>
<td>1.02</td>
<td>(0.97, 1.06)</td>
<td>1.11</td>
<td>(1.06, 1.16)</td>
<td>1.31</td>
<td>(1.24, 1.38)</td>
</tr>
<tr>
<td>40%</td>
<td>Stage I</td>
<td>3.30</td>
<td>(3.04, 3.57)</td>
<td>2.70</td>
<td>(2.45, 2.96)</td>
<td>2.14</td>
<td>(1.91, 2.38)</td>
</tr>
<tr>
<td></td>
<td>Stage II</td>
<td>1.93</td>
<td>(1.85, 2.00)</td>
<td>2.13</td>
<td>(2.02, 2.25)</td>
<td>2.19</td>
<td>(1.99, 2.40)</td>
</tr>
<tr>
<td></td>
<td>Stage III</td>
<td>1.01</td>
<td>(0.97, 1.06)</td>
<td>1.11</td>
<td>(1.06, 1.16)</td>
<td>1.31</td>
<td>(1.24, 1.38)</td>
</tr>
</tbody>
</table>

respectively, indicating that the ‘city lockdown’ of the Wuhan was effective. After the ‘city lockdown’, a series of strict prevention and control measures have been taken nationwide, such as restriction on outdoor activities or gathering, in-home work scheme. Thus, in stage III, the reproduction number decreased further to 0.98, 1.14 and 1.41 respectively.

Discussion

At present, the reproduction number is estimated at 0.98 (95% CI: 0.96–1.03). In Hubei province, the epidemic situation turns better, with reproduction number 1.14 (95% CI: 1.10–1.19). Although it is still larger than 1 significantly, the downward trend is stable and is likely to be less than 1 shortly in the future. However, the reproduction number in Wuhan is estimated at 1.41 (95% CI: 1.35–1.48) and thus, enhancement in the control measures is supposed to be maintained.

Since January 27, a new coronavirus nucleic acid detection kit has been launched around the country, which greatly improves the detection rate of the disease. However, severe under-reporting phenomena caused by the imbalance regional detection level existed before the appearance of the new coronavirus nucleic acid detection kit (11,22). It is difficult to estimate the specific value of under-reporting rate. According to (16), under-reporting has a very small influence on the reproduction number calculated by this method. In this paper, we assumed that the reporting rate was 20%, 30% and 40% respectively. After correction, we recalculated the average daily reproduction numbers (Table 2) of three stages of China, Hubei and Wuhan.

From Table 2, the average reproduction number of the stage I was lower than that of the original data which was not corrected. Therefore, the actual average reproduction number in the stage I might not be as large as that in Table 1, and the risk of disease had been overestimated. In the stage II and stage III, the rate of under-reporting had little effect on their reproduction numbers.

In order to protect the lives and health of the people, guarantee the prevention and control of the epidemic and ensure transportation of production and living materials, etc., from February 17 to the end of the epidemic prevention and control work, all vehicles are exempt from express away toll according to notice of Ministry of Transport of China (23) to encourage travelling by private cars. Meanwhile, the passenger number of long-distance buses and waterway passenger liners should be controlled under 50% of corresponding passenger load factors and passengers are to be guided to maintain a certain distance (24). This might reduce the possibility of a second outbreak.

In fact, the number of new cases per day in other provinces (exclude Hubei) has declined continually for 13 days, and the epidemic is expected to end in the near future. The research data in this paper is as of February 16, and in the future, the transmission of the disease is supposed to be considered again and always be considered when calculating a new reproduction number.

Conclusions

We estimated the time-varying reproduction numbers
of China, Hubei and Wuhan. The results indicate the improvement and effectiveness of the COVID-19 control measures, although the population in Wuhan is still at a higher risk of infection. As such, we recommend to maintain the enhancement in the public health control efforts in combating the ongoing COVID-19 epidemic.

Acknowledgments

Funding: This work was supported in part by the Natural Science Foundation of China (11961071, 81673275), PhD research startup foundation of Xinjiang Medical University (2018-12), the National S&T Major Project Foundation of China (2018ZX10715002–004–002, 2018ZX10713001–001), program for Tianshan Innovative Research Team of Xinjiang Uygur Autonomous Region, China (2020D14020).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-1944). DH reports grants from Alibaba (China) - HKPU collaborative research project, during the conduct of the study. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The data of the confirmed COVID-19 cases were collected via the National Health Commission of the People’s Republic of China, and thus neither ethical approval nor individual consent was not applicable.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial repilation and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


## Table S1 Average daily reproduction numbers for cases of COVID-19 in different stages in China, Hubei and Wuhan with SI at 7.5 days

<table>
<thead>
<tr>
<th>Period</th>
<th>China</th>
<th></th>
<th>Hubei</th>
<th></th>
<th>Wuhan</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>95% CI</td>
<td>R</td>
<td>95% CI</td>
<td>R</td>
<td>95% CI</td>
</tr>
<tr>
<td>Stage I</td>
<td>8.29</td>
<td>(7.74, 8.86)</td>
<td>7.10</td>
<td>(6.54, 7.66)</td>
<td>4.85</td>
<td>(4.37, 5.32)</td>
</tr>
<tr>
<td>Stage II</td>
<td>3.26</td>
<td>(3.13, 3.36)</td>
<td>3.76</td>
<td>(3.57, 3.96)</td>
<td>3.71</td>
<td>(3.39, 4.03)</td>
</tr>
<tr>
<td>Stage III</td>
<td>0.97</td>
<td>(0.89, 1.07)</td>
<td>1.12</td>
<td>(1.02, 1.22)</td>
<td>1.44</td>
<td>(1.31, 1.59)</td>
</tr>
</tbody>
</table>

## Table S2 The effect of under-report rate in the early outbreak on the average daily reproduction number with different stages in China, Hubei and Wuhan with SI at 7.5 days

<table>
<thead>
<tr>
<th>Under-report rate</th>
<th>Period</th>
<th>China</th>
<th></th>
<th>Hubei</th>
<th></th>
<th>Wuhan</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>95% CI</td>
<td>R</td>
<td>95% CI</td>
<td>R</td>
<td>95% CI</td>
</tr>
<tr>
<td>20%</td>
<td>Stage I</td>
<td>7.68</td>
<td>(7.21, 8.16)</td>
<td>6.39</td>
<td>(5.92, 6.87)</td>
<td>4.28</td>
<td>(3.89, 4.69)</td>
</tr>
<tr>
<td></td>
<td>Stage II</td>
<td>2.95</td>
<td>(2.84, 3.05)</td>
<td>3.44</td>
<td>(3.27, 3.60)</td>
<td>3.47</td>
<td>(3.19, 3.75)</td>
</tr>
<tr>
<td></td>
<td>Stage III</td>
<td>0.97</td>
<td>(0.89, 1.06)</td>
<td>1.11</td>
<td>(1.02, 1.22)</td>
<td>1.44</td>
<td>(1.31, 1.58)</td>
</tr>
<tr>
<td>30%</td>
<td>Stage I</td>
<td>7.36</td>
<td>(6.92, 7.80)</td>
<td>6.02</td>
<td>(5.59, 6.45)</td>
<td>3.99</td>
<td>(3.64, 4.35)</td>
</tr>
<tr>
<td></td>
<td>Stage II</td>
<td>2.77</td>
<td>(2.68, 2.87)</td>
<td>3.25</td>
<td>(3.10, 3.40)</td>
<td>3.33</td>
<td>(3.07, 3.60)</td>
</tr>
<tr>
<td></td>
<td>Stage III</td>
<td>0.96</td>
<td>(0.88, 1.06)</td>
<td>1.11</td>
<td>(1.02, 1.22)</td>
<td>1.44</td>
<td>(1.31, 1.58)</td>
</tr>
<tr>
<td>40%</td>
<td>Stage I</td>
<td>7.03</td>
<td>(6.64, 7.43)</td>
<td>5.64</td>
<td>(5.26, 6.03)</td>
<td>3.69</td>
<td>(3.37, 4.01)</td>
</tr>
<tr>
<td></td>
<td>Stage II</td>
<td>2.59</td>
<td>(2.50, 2.68)</td>
<td>3.05</td>
<td>(2.91, 3.19)</td>
<td>3.18</td>
<td>(2.94, 3.43)</td>
</tr>
<tr>
<td></td>
<td>Stage III</td>
<td>0.96</td>
<td>(0.88, 1.05)</td>
<td>1.11</td>
<td>(1.01, 1.21)</td>
<td>1.44</td>
<td>(1.30, 1.58)</td>
</tr>
</tbody>
</table>