medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Preliminary estimating the reproduction number of the coronavirus disease (COVID-19) outbreak in Republic of Korea from 31 January to 1 March 2020

Zian Zhuang[1], Shi Zhao[2,3], Qianying Lin[4], Peihua Cao[5], Yijun Lou[1], Lin Yang[6,*], Shu Yang[7] and Daihai He[1,*]

1 Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong SAR, China

2 JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China

3 Shenzhen Research Institute of Chinese University of Hong Kong, Shenzhen, China

4 Michigan Institute for Data Science, University of Michigan, Ann Arbor, Michigan, USA

5 Clinical Research Centre, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China

6 School of Nursing, Hong Kong Polytechnic University, Hong Kong, China

7 College of Medical Information Engineering, Chengdu University of Traditional Chinese Medicine, Chengdu, China

* Correspondence to: l.yang@polyu.edu.hk (LY), and daihai.he@polyu.edu.hk (DH)

medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract:

The novel coronavirus disease 2019 (COVID-19) outbreak in Republic of Korea has caused 3736 cases and 18 deaths by 1 March 2020. We modeled the transmission process in Republic of Korea with a stochastic model and estimated the basic reproduction number R_0 as 2.6 (95%CI: 2.3-2.9) and 3.2 (95%CI: 2.9-3.5), under the assumption that the exponential growth starting 31 January and 5 February, 2020, respectively. Estimates of dispersion term (k) were larger than 10 significantly, which implies few super-spreading events..

Keywords:

coronavirus disease 2019; the basic reproduction number; dispersion term; Republic of Korea

medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

1. Introduction

The coronavirus disease 2019 (COVID-19) first emerged in Wuhan, China in the end of 2019 and spread to more than 60 foreign countries as of 1 March 2020 [1]. On 20 January 2020, the first imported COVID-19 case was detected in Republic of Korea, the epidemic curve appeared steadily until 15 February. In the second half of February, the number of reported cases increased rapidly with more than 1200 cases a week. As of 1 March 2020, there were 3736 cases confirmed including 18 deaths [1]. To date, there are 7169 confirmed cases in the world except China, which means that nearly half of cases outside China are from Republic of Korea [1].

In this study, we modelled the early outbreak of COVID-19 in Republic of Korea and estimated the basic reproduction number under different transmission starting date.

2. Methods

We collect time series of reported COVID-19 cases in Republic of Korea from 20 January to 1 March 2020 [1].

Following [2,3], we assumed that number of secondary cases associated with a primary COVID-19 case follows a negative binomial (NB) distribution, with mean R_0 and dispersion parameter k [3]. Here, the R_0 is the basic reproduction number of COVID-19. The k measures the likelihood of the superspreading events, and a larger k means the superspreading is less likely to occur. The onset date of each secondary case is the summation of the onset date of the primary case (t) plus the serial interval (SI). In this work, the SI was assumed as a Gamma distribution with a 4.5-day mean and a 3.1-day standard deviation (SD) [4-7]. The transmission process was simulated stochastically.

Since the number of early reported cases are stable, which appears no sign of the outbreak in January 2020, we consider the self-sustaining transmission might start since February 2020. Hence, we simulated the transmission starting on

- 31 January 2020, and
- 5 February 2020

medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

with one seed infection. The R_0 and k were estimated by the maximum likelihood estimates approach that fit the reported cases with Poisson-distributed likelihood framework as follows.

$$l(\lambda) = \log \mathcal{L}(\lambda) = \sum_{i=1}^{T} \log f(n_t; R_0, k). \quad (1)$$

Here, the $l(\cdot)$ is the overall log-likelihood and T is the total number of days since the start of transmission. The n_t represents number of cases reported on t-th day. We calculated 95% confidence intervals (95%CI) by using the profile likelihood estimation approach determined by a Chi-square quantile.

3. Results and discussion

In Table 1, we estimated the R_0 at 2.6 (95% CI: 2.3−2.9) and 3.2 (95% CI: 2.9−3.5) with the transmission starting date on 31 January and 5 February 2020 respectively. Estimates of dispersion term (k) were larger than 10 significantly, which was consistent with [3] and suggested the unlikelihood of superspreading events.

Fig 1 showed that the fitting results matched the observed number of cases well. Our estimated R_0 in Republic of Korea is relatively higher than that in Shim *et al.*'s [8], which set the transmission starting date on 20 January 2020 and the onset date data were considered. Here we use laboratory confirmation date data which covered both symptomatic and asymptomatic cases. In addition, the R_0 estimate was largely consistent with those estimates based on the epidemic curve in China [9,10]. If there were evidences suggesting that the exponential growth started earlier, we note that the R_0 estimate would decrease. Public activities in the late date and cold weather in our study period could have speed up the transmission which explains the higher estimates.

The ongoing COVID-19 outbreak in Republic of Korea could be amplified due to large-scale gathering activities [11]. Without public health control or self-protective measures, the epidemic was likely to grow in a relatively large rate. The R_0 estimate and the rapid growth of epidemic curves both indicates the disease transmissibility. Thus, control measures as well as self-administered protective actions are crucial to reduce the transmissibility of COVID-19 and thus mitigate the outbreak size and

medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

prevent for further burden. Given the superspreading is unlikely to occur, we can be confident to control the COVID-19 outbreak by reducing the reproduction number to below unity.

Table 1. The summary of the basic reproduction number and dispersion parameter estimates under different starting transmission date.

transmission starting date	R_0	dispersion
31 January	2.6 (2.3, 2.9)	> 10
5 February	3.2 (2.9, 3.5)	> 10

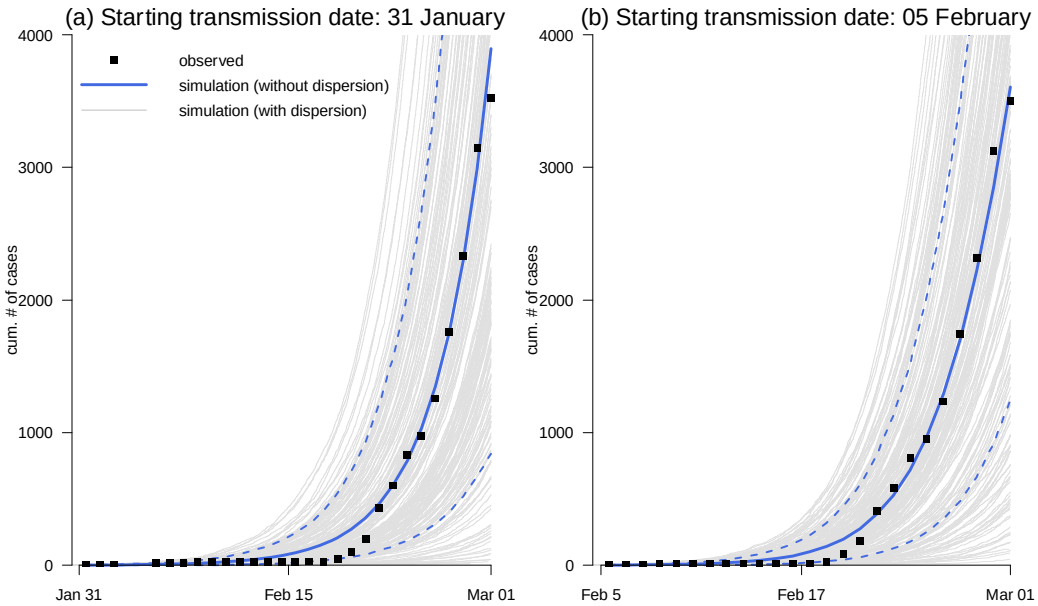

Figure 1. The observed (dots) and fitted (curves) number of COVID-19 cases in Republic of Korea. Panels (a) and (b) show the results with the transmission starting date on 31 January and 5 February 2020 respectively. In both panels, the grey curves are 1000 simulations, the blue bold curve is the simulation median, and the blue dashed curves are the 95%CI.

medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Coronavirus disease 2019 (COVID-19) Situation Report – 41. Availabe online: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/ (accessed on 1 March 2020).
2. Althaus, C.L. Ebola superspreading. *The Lancet Infectious Diseases* **2015**, *15*, 507-508, doi:10.1016/s1473-3099(15)70135-0.
3. Riou, J.; Althaus, C.L. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill* **2020**, *25*, 2000058, doi:10.2807/1560-7917.ES.2020.25.4.2000058.
4. Zhao, S.; Gao, D.; Zhuang, Z.; Chong, M.; Cai, Y.; Ran, J.; Cao, P.; Wang, K.; Lou, Y.; Wang, W., et al. Estimating the serial interval of the novel coronavirus disease (COVID-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020. *medRxiv* **2020**, 10.1101/2020.02.21.20026559, 2020.2002.2021.20026559, doi:10.1101/2020.02.21.20026559.
5. Du, Z.; Wang, L.; Xu, X.; Wu, Y.; Cowling, B.J.; Meyers, L.A. The serial interval of COVID-19 from publicly reported confirmed cases. *medRxiv* **2020**, 10.1101/2020.02.19.20025452, 2020.2002.2019.20025452, doi:10.1101/2020.02.19.20025452.
6. Nishiura H, Linton NM, Akhmetzhanov AR: Serial interval of novel coronavirus (2019-nCoV) infections. medRxiv 2020:2020.2002.2003.20019497.
7. Jung SM, Akhmetzhanov AR, Hayashi K, Linton NM, Yang Y, Yuan B, Kobayashi T, Kinoshita R, Nishiura H: Real-Time Estimation of the Risk of Death from Novel Coronavirus (COVID-19) Infection: Inference Using Exported Cases. J Clin Med 2020, 9(2).
8. Shim, E.; Tariq, A.; Choi, W.; Lee, Y.; Chowell, G. Transmission potential of COVID-19 in Republic of Korea. *medRxiv* **2020**, 10.1101/2020.02.27.20028829, 2020.2002.2027.20028829, doi:10.1101/2020.02.27.20028829.
9. Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, Lou Y, Gao D, Yang L, He D, Wang MH. 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. *International Journal of Infectious Diseases*. **2020**, 92;214-217

medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

10. Wu JT, Leung K, Leung GM: Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet 2020.
11. Half of Republic of Korea's coronavirus cases are linked to a controversial religious organization. Availabe online: https://www.vox.com/2020/2/22/21148477/coronavirus-south-korea-shincheonji-church-jesus (accessed on 1 March 2020).

Declarations

Ethics approval and consent to participate

The ethical approval or individual consent was not applicable.

Availability of data and materials

All data and materials used in this work were publicly available.

Consent for publication

Not applicable.

Funding

DH was supported by General Research Fund (15205119) of Research Grants Council of Hong Kong and an Alibaba-Hong Kong Polytechnic University Collaborative Research project.

Acknowledgements

None.

Disclaimer

The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Competing Interests

DH was supported by an Alibaba-Hong Kong Polytechnic University Collaborative Research project. Other authors declare no competing interests.

Authors' Contributions

medRxiv preprint doi: https://doi.org/10.1101/2020.03.02.20030312. The copyright holder for this preprint **(which was not peer-reviewed)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

All authors conceived the study, carried out the analysis, discussed the results, drafted the first manuscript, critically read and revised the manuscript, and gave final approval for publication.