Preliminary estimating the reproduction number of the coronavirus disease (COVID-19) outbreak in Republic of Korea and Italy by 5 March 2020

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

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: daihai.he@polyu.edu.hk (DH), l.yang@polyu.edu.hk (LY), and xiaolichengdu@163.com

Abstract:

The novel coronavirus disease 2019 (COVID-19) outbreak and Italy has caused 6088 cases and 41 deaths in Republic of Korea and 3144 cases and 107 death in Italy by 5 March 2020. We modeled the transmission process in Republic of Korea and Italy with a stochastic model and estimated the basic reproduction number R_0 as 2.6 (95% CI: 2.3-2.9) or 3.2 (95% CI: 2.9-3.5) in Republic of Korea, under the assumption that the exponential growth starting on 31 January or 5 February 2020, and 2.6 (95% CI: 2.3-2.9) or 3.3 (95% CI: 3.0-3.6) in Italy, under the assumption that the exponential growth starting on 5 February or 10 February 2020. Estimates of dispersion term (k) were 10 (95% CI: 5-56) or 22 (95% CI: 8-61) in Republic of Korea, and 13 (95% CI: 5-61) or 37 (95% CI: 13-61) in Italy, and all of which imply few super-spreading events.

Keywords:

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

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, and 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 5 March 2020, there were 6088 cases confirmed including 41 deaths [1]. In Italy, the first case was reported on 6 February 2020 and the epidemic curve was steadily by 21 February. Then the number of reported cases soared rapidly, reaching 3142 reported cases and 107 deaths until 5 March 2020 [1]. To date, there are 14768 confirmed cases in the world except China, which means that more than 60 per cent of cases outside China are from Republic of Korea and Italy [1].
In this study, we modelled the early outbreak of COVID-19 in Republic of Korea and Italy to estimate the basic reproduction number under different exponential growth starting date.

2. Methods

We collect time series of reported COVID-19 cases in Republic of Korea from 20 January to 1 March 2020 and cases in Italy from 5 February to 5 March 2020. Following [2,3], we assumed that number of secondary cases associated with a primary COVID-19 case follows a negative binomial (NB) distribution, with means 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 occurrence of super-spreading events (or other factors) which could vary the growth rate. If k is smaller than 1, these data indicate occurrence of super-spreading events (larger proportion of 'super-spreaders' or 'dead-ends' in infected individuals) or some other factors which could vary the growth rate; otherwise if k is larger than 1, these data do not indicate occurrence of super-spreading events [4].

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) [5-8]. 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 sustaining exponential growth might start since February 2020. Hence, we simulated the exponential growth starting on

- 31 January and 5 February 2020 for Republic of Korea
- 5 February and 10 February 2020 for Italy

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 exponential growth. 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 in Republic of Korea 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, and the R_0 in Italy at 2.6 (95% CI: 2.3-2.9) and 3.3 (95% CI: 3.0-3.6), with the transmission starting date on 5 February and 10 February 2020. Estimates of dispersion term (k) for Republic of Korea were 10 (95%CI: 5-56) and 22 (95%CI: 8-61), under the assumption that the exponential growth starting on 31 January and 5 February 2020, and for Italy were 13 (95% CI: 5-61) and 37 (95% CI: 13-61), under the assumption that the exponential growth starting on 5 February and 10 February 2020, all of which were 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 [9], which set the exponential growth 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 estimates from two countries were largely consistent with those estimates based on the epidemic curve in China [10,11]. If there were evidences suggesting that the exponential growth started earlier, we note that the R_0 estimates 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 and Italy could be amplified due to large-scale gathering activities [12,13]. Without public health control or self-protective measures, the epidemic was likely to grow in a relatively large rate. The R_0 estimates 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 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 exponential growth starting date.

Country	starting date	R_0 (95% CI)	Dispersion (95% CI)
Republic of Korea	31 January	2.6 (2.3, 2.9)	10 (5, 56)
Republic of Korea	5 February	3.2 (2.9, 3.5)	22 (8, 61)
Italy	5 February	2.6 (2.3, 2.9)	13 (5, 61)
Italy	10 February	3.3 (3.0, 3.6)	37 (13, 61)

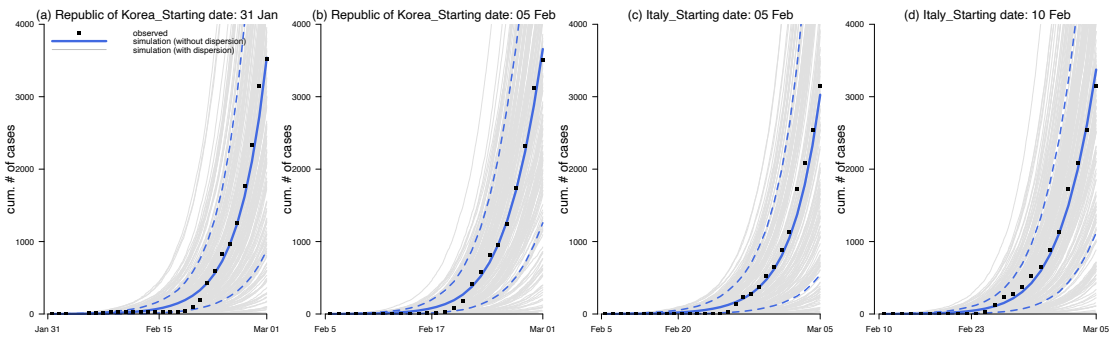

Figure 1. The observed (dots) and fitted (curves) number of COVID-19 cases in Republic of Korea and Italy. Panels (a) and (b) show the results in Republic of Korea with the exponential growth starting date on 31 January and 5 February 2020 respectively. Panels (c) and (d) show the results in Italy with the exponential growth starting date on 5 February and 10 February 2020 respectively. In all panels, the grey curves are 1000 simulations, the blue bold curve is the simulation median, and the blue dashed curves are the 95%CI.

References

1. Wu, T.; Ge, X.; Yu, G.; Hu, E. Open-source analytics tools for studying the COVID-19 coronavirus outbreak. *medRxiv* **2020**, 10.1101/2020.02.25.20027433, 2020.2002.2025.20027433, doi:10.1101/2020.02.25.20027433.
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. Hartfield, M.; Alizon, S. Introducing the outbreak threshold in epidemiology. *PLoS Pathog* **2013**, *9*, e1003277, doi:10.1371/journal.ppat.1003277.
5. 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.
6. 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.

7. Nishiura, H.; Linton, N.M.; Akhmetzhanov, A.R. Serial interval of novel coronavirus (2019-nCoV) infections. *medRxiv* **2020**, 10.1101/2020.02.03.20019497, 2020.2002.2003.20019497, doi:10.1101/2020.02.03.20019497.
8. Jung, S.M.; Akhmetzhanov, A.R.; Hayashi, K.; Linton, N.M.; 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*, doi:10.3390/jcm9020523.
9. 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.
10. Zhao, S.; Lin, Q.; Ran, J.; Musa, S.S.; Yang, G.; Wang, W.; Lou, Y.; Gao, D.; Yang, L.; He, D., et al. 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, doi:10.1016/j.ijid.2020.01.050.
11. Wu, J.T.; Leung, K.; Leung, G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet* **2020**, *395*, 689-697, doi:10.1016/S0140-6736(20)30260-9.
12. 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 5 March 2020).
13. Italian town shows zest for history with 'Battle of the Oranges' food fight. Availabe online: https://abcnews.go.com/International/italian-town-shows-zest-history-battle-oranges-food/story?id=61452172 (accessed on 5 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

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.