Short Communication

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

Zian Zhuang[a], Shi Zhao[b,c], Qianying Lin[d], Peihua Cao[e], Yijun Lou[a], Lin Yang[f,*], Shu Yang[g], Daihai He[a,*], Li Xiao[g,*]

[a] Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China
[b] JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China
[c] Shenzhen Research Institute of Chinese, University of Hong Kong, Shenzhen, China
[d] Michigan Institute for Data Science, University of Michigan, Ann Arbor, MI, USA
[e] Clinical Research Centre, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China
[f] School of Nursing, Hong Kong Polytechnic University, Hong Kong, China
[g] College of Medical Information Engineering, Chengdu University of Traditional Chinese Medicine, Chengdu, China

A B S T R A C T

The novel coronavirus disease 2019 (COVID-19) outbreak has caused 6088 cases and 41 deaths in Republic of Korea, and 3144 cases and 107 death in Italy by 5 March 2020, respectively. We modelled the transmission process in the 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 the 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, respectively.

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 (Wu et al., 2020). 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 1,200 cases a week. As of 5 March 2020, there were 6,088 cases confirmed including 41 deaths (Wu et al., 2020). 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 3,142 reported cases and 107 deaths until 5 March 2020 (Wu et al., 2020). To date, there are 14,768 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.

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.

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 (Althaus, 2015), 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 a dispersion parameter k. Here, the R_0 is the basic reproduction number of COVID-19.

The onset date of each secondary case is the summation of the onset date of the primary case (t) plus the one serial interval (SI). In this work, the SI was assumed to follow a Gamma distribution with a 4.5-day mean and a 3.1-day standard deviation (SD) (Zhao et al., 2020a; Du et al., 2020; Nishiura et al., 2020). The transmission process was simulated stochastically.

Since the number of early reported cases remained at a low level, and there was no sign of outbreak in January 2020, we assumed that the sustaining exponential growth might have

Table 1

The summary of the basic reproduction number estimates under different exponential growth starting date.

Country	Starting date	R_0 (95% CI)
Republic of Korea	31 January	2.6 (2.3, 2.9)
Republic of Korea	5 February	3.2 (2.9, 3.5)
Italy	5 February	2.6 (2.3, 2.9)
Italy	10 February	3.3 (3.0, 3.6)

started since February 2020. Hence, we simulated the start of an exponential growth on the following dates with one seed infection:

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

The value of R_0 was estimated by the maximum likelihood estimation approach that fitted the reported cases with the Poisson-distributed likelihood framework as follows:

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

Here, $l(\cdot)$ is the overall log-likelihood, and T is the total number of days since the start of the 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.

Results and discussion

In Table 1, we estimated that the R_0 of COVID-19 in the Republic of Korea was 2.6 (95% CI: 2.3–2.9) and 3.2 (95% CI: 2.9–3.5) when the transmission was assumed to start on 31 January or 5 February 2020, respectively. The R_0 in Italy was 2.6 (95% CI: 2.3–2.9) and 3.3

(95% CI: 3.0–3.6), under the assumptions that the transmission started on 5 February or 10 February 2020, respectively.

Figure 1 shows that the model estimates well matched the observed number of cases. Our estimated R_0 in thevRepublic of Korea is relatively higher than that in Shim *et al.*'s (Shim et al., 2020), which set the exponential growth starting date on 20 January 2020 and fit the model according to the onset dates. Here, we used the daily data by laboratory confirmation dates, in order to cover both symptomatic and asymptomatic cases. In addition, the R_0 estimates from two countries were largely consistent with those estimates in Wuhan, China (Zhao et al., 2020b). If the exponential growth started earlier, we found that the R_0 estimates would decrease. Public activities and cold weather in our study period could have escalated the transmission which might help explain the higher estimates.

The ongoing COVID-19 outbreak in the Republic of Korea and Italy could be amplified by large-scale gathering activities (VOX, 2020; ABC NEWS, 2020). Without public health control or self-protective measures, the epidemic was likely to grow in a relatively fast rate. Thus, control measures adopted by the governments and the public are crucial to reduce the transmissibility of COVID-19 and thereby mitigating the outbreak size to minimize health burden.

In Supplementary Fig. S1, we show the daily cumulative COVID-19 cases, from the dates exceeding 100 cumulative cases till April 11, 2020, for 16 countries. The Republic of Korea and Italy would rank very high among these countries if no control measures were taken. In Supplementary Fig S2, we show the daily reported cases and their instantaneous reproductive number (R_e) estimated using the **R** package 'EpiEstim' (Cori et al., 2013) with the aforementioned SI distribution. Their initial instantaneous reproductive number can be regarded as the basic reproductive number. The pink bold curve shows the estimates of R_e with 95% confidence intervals. The epidemic will fade out when the R_e is below 1. In principle, only local cases should be used, which is held for all countries except China, where the second wave is largely due to imported cases.
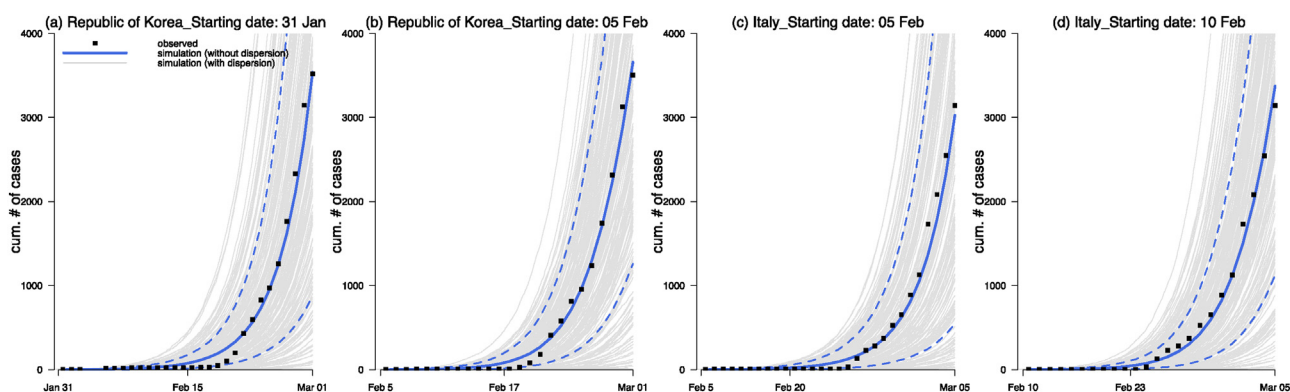

Figure 1. The observed (dots) and fitted (curves) number of COVID-19 cases in the Republic of Korea and Italy. Panels (a) and (b) show the results in the Republic of Korea with the exponential growth startingon 31 January and 5 February 2020 respectively. Panels (c) and (d) show the results in Italy with the exponential growth starting 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.

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

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 (China)-Hong Kong Polytechnic University Collaborative Research project. Other authors declared 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.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at https://doi.org/10.1016/j.ijid.2020.04.044.

References

Italian town shows zest for history with 'Battle of the Oranges' food fight. Available online: https://abcnews.go.com/International/italian-town-shows-zest-history-battle-oranges-food/story?id=61452172 [accessed 05.03.20].

Althaus CL. Ebola superspreading. Lancet Infect Dis 2015;15:507–8, doi:http://dx.doi.org/10.1016/s1473-3099(15)70135-0.

Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol 2013;178(9):1505–12.

Du Z, Wang L, Xu X, Wu Y, Cowling BJ, Meyers LA. The serial interval of COVID-19 from publicly reported confirmed cases. medRxiv 2020;, doi:http://dx.doi.org/10.1101/2020.02.19.20025452.

Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (2019-nCoV) infections. medRxiv 2020;, doi:http://dx.doi.org/10.1101/2020.02.03.20019497.

Shim E, Tariq A, Choi W, Lee Y, Chowell G. Transmission potential of COVID-19 in Republic of Korea. medRxiv 2020;, doi:http://dx.doi.org/10.1101/2020.02.27.20028829.

Half of Republic of Korea's coronavirus cases are linked to a controversial religious organization. Available online: https://www.vox.com/2020/2/22/21148477/coronavirus-south-korea-shincheonji-church-jesus [accessed 05.03.20].

Wu T, Ge X, Yu G, Hu E. Open-source analytics tools for studying the COVID-19 coronavirus outbreak. medRxiv 2020;, doi:http://dx.doi.org/10.1101/2020.02.25.20027433.

Zhao S, Gao D, Zhuang Z, Chong M, Cai Y, Ran J, 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. medRxiv 2020a;15:, doi:http://dx.doi.org/10.1101/2020.02.21.20026559.

Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, 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. Int J Infect Dis 2020b;92:214–7, doi:http://dx.doi.org/10.1016/j.ijid.2020.01.050.