

# Temporal Changes in the Risk of Superspreading Events of Coronavirus Disease 2019

Jun-Sik Lim,<sup>1,2</sup> Eunbi Noh,<sup>1,3</sup> Eunha Shim,<sup>4</sup> and Sukhyun Ryu<sup>3,6</sup>

<sup>1</sup>Graduate School of Public Health, Seoul National University, Seoul, South Korea, <sup>2</sup>College of Veterinary Medicine and Institute of Veterinary Science, Kangwon National University, Chuncheon, South Korea, <sup>3</sup>Department of Preventive Medicine, Konyang University College of Medicine, Daejeon, South Korea, and <sup>4</sup>Department of Mathematics, Soongsil University, Seoul, South Korea

To identify the temporal change in the possible risk of superspreading events, we estimated the overdispersion parameter in 2 different periods of the coronavirus disease 2019 pandemic. We determined that the possible risk of superspreading events was reduced 90% during the second epidemic period in South Korea.

**Keywords.** coronavirus; infectiousness; overdispersion; superspreading; transmission.

Individual-based interventions, including active contact tracing and isolation, interrupt the transmission chain and reduce the infectiousness of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1, 2]. To identify changes in the transmissibility of SARS-CoV-2, the daily effective reproductive number has been estimated. However, this estimate has limited interpretability, as it does not reflect the individual variation of infectiousness [3], which is known as overdispersion in the distribution of the number of secondary transmissions and is the deterministic factor of superspreading events (SSEs) [4].

The  $k$  is assumed to be a fixed characteristic of infection, and there is lack of study demonstrating its temporal changes on different epidemic periods. Here, we examined contact tracing data in South Korea to assess temporal changes in the individual variation of the risk of SSEs on SARS-CoV-2 transmission.

## METHODS

We retrieved publicly available data on coronavirus disease 2019 (COVID-19) cases from local public health authorities in South

Korea. The authorities provided daily updates of all COVID-19 cases confirmed by consistent real-time reverse-transcription polymerase chain reaction (PCR). The data comprised the date of illness onset, date of laboratory diagnosis, and infection source. We excluded the Daegu-Gyeongsangbuk region, where data were not publicly reported.

Based on the epidemic curve of 2 epidemic waves in South Korea [5], we divided the study period into 2, based on the date of symptom onset of the infector in pairs: period 1 (19 January–19 April 2020) and period 2 (20 April–16 October 2020). Furthermore, based on the infection source, we built infectee–infector pairs to estimate the offspring distribution.

Offspring distribution was estimated from the infectee–infector pairs to demonstrate the overdispersion of individual heterogeneity in transmission for each epidemic period. Using the observed distribution of the secondary case, we fitted the negative binomial offspring distribution (Supplementary Figure). This distribution was parameterized with the basic reproductive number ( $R_0$ ) and dispersion parameter ( $k$ ), which indicate the mean number of secondary cases generated by an index case and the individual level of heterogeneity in transmission, respectively [6, 7]. A smaller  $k$  indicates a longer tail of the negative binomial distribution, which means the individual level of variation in secondary cases is higher, and hence the epidemic includes a higher likelihood of SSEs [3]. This distribution has been widely used to identify the individual level of variation in infectiousness [6–8]. Furthermore, we estimated the expected proportion of cases responsible for 80% of the total secondary transmission and the probability of SSEs using the estimated  $R_i$  and  $k_i$ , which was taken from previous studies [4, 8], and the probabilities that an index case of SARS-CoV-2 infection results in a cluster of 10 cases or more were estimated (Supplementary Appendix). We used the *runjags* package for Markov chain Monte Carlo simulation. All statistical analyses were performed in R version 3.6.1 (R Foundation for Statistical Computing).

## RESULTS

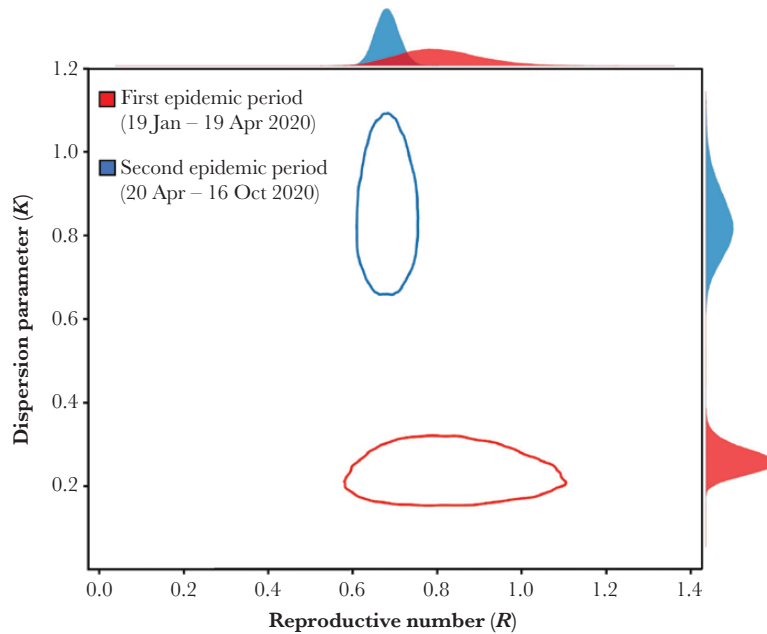
In this study, we identified 675 transmission pairs (104 for period 1 and 571 for period 2) with 1070 sporadic cases not linked with other cases (240 for period 1 and 830 for period 2). Our estimates suggested that the  $k$  in period 1 (0.23 [95% credible interval {CrI}, .15–.28]) was lower than period 2 (0.85 [95% CrI, .70–1.05]) (Figure 1). Furthermore, the expected proportion of cases responsible for 80% of secondary transmission was 14.70% (95% CrI, 12.21%–17.36%) in period 1 and 25.72% (95% CrI, 24.02%–27.41%) in period 2. The

Received 20 March 2021; editorial decision 24 June 2021; accepted 29 June 2021.

Correspondence: Sukhyun Ryu, MD, PhD, Department of Preventive Medicine, Konyang University College of Medicine, R707, Myungok-Euihak Gwan, 158, Gwanjeodong-ro, Seogu, Daejeon, South Korea ([gentryu@onehealth.or.kr](mailto:gentryu@onehealth.or.kr)).

### Open Forum Infectious Diseases® 2021

© The Author(s) 2021. Published by Oxford University Press on behalf of Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact [permissions@oup.com](mailto:permissions@oup.com) <https://doi.org/10.1093/ofid/ofab350>

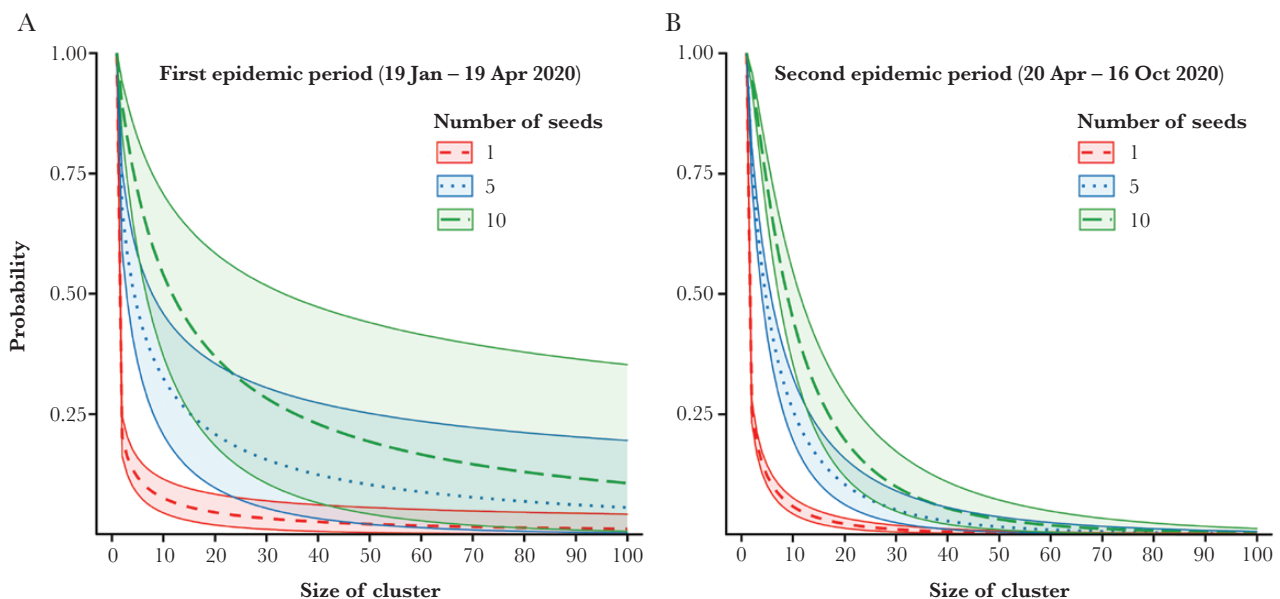


**Figure 1.** Joint estimates of overdispersion parameter and reproductive number of coronavirus disease 2019. The red and blue lines indicate the 95% credible intervals of the estimated overdispersion parameter ( $k$ ) in the first epidemic period (19 January–19 April 2020) and second epidemic period (20 April–16 October 2020), respectively. The posterior distributions of each parameter were plotted in the outer margin of each axis.

probability of SSEs was 2.43% (95% CrI, 1.39%–4.04%) in period 1 and 0.24% (95% CrI, .14%–.40%) in period 2. The probability that an index case of SARS-CoV-2 infection results in a cluster of 10 cases or more was 7.46% (95% CrI, 4.47%–11.65%) in period 1 and 5.74% (95% CrI, 4.24%–7.56%) in period 2 (Figure 2).

## DISCUSSION

Our findings suggest that the individual variation of SARS-CoV-2 infectiousness is highly overdispersed, which is corroborated by previous reports that  $k$  was estimated to be 0.43 in Hong Kong [8]. Furthermore, our results showed that  $k$  was larger in the latter epidemic period, which is consistent with a previous study that the serial



**Figure 2.** The expected probabilities that the different numbers of index cases generate clusters of given size during (A) the first epidemic period (19 Jan–19 April 2020), and (B) the second epidemic period (20 Apr–16 Oct 2020). Dashed, dotted, and long-dashed lines indicate the median estimates of the expected probabilities that 1, 5, and 10 infectors generate clusters of a given size, respectively. The green, red, and blue areas covered by solid lines indicate 95% credible intervals.

interval was reduced over time as enhanced nonpharmaceutical intervention was implemented in the community [2]. This reduction suggested that the risk of SSEs was lower in the latter period of the epidemic. This may be affected by improved contact tracing with the digital QR code entry logs at public facilities, which began 10 June 2020, to quickly identify the possible COVID-19 case [9] and enhanced social distancing in publicly used facilities (a ban on religious gatherings began 19 August 2020) during period 2 [10]. Other factors including the public's behavior changes, derived by improved risk awareness of COVID-19 and climatological factors, could also influence the transmission dynamics of SARS-CoV-2 and the changes of the risk of SSEs.

The South Korean public health authority has modified the level of nationwide nonpharmaceutical interventions, primarily using data on the mobility of people, the number of COVID-19 cases, the number of clusters, and the effective reproductive number [11]. Because few cases could account for many COVID-19 cases, monitoring the overdispersion parameter would help to provide relevant information on transmission dynamics of SARS-CoV-2 to the public health authority.

There are several limitations to our study. First, we used a case report from public health authorities that is not free from underreporting of cases and could bias our results. Therefore, our result of the possible risk of SSEs is likely to be a lower bound estimate. Second, we have not included cases from the Daegu-Gyeongsangbuk region, which originated primarily from the Shincheonji religious group. A previous report demonstrated that there was a delay of several days in detecting cases that could affect the size of the cluster [12]. Therefore, this group may not reflect the typical characteristics of transmission. Third, we have not included the latter period of the COVID-19 pandemic in which the number of COVID-19 cases was larger than previous 2 periods in the present study. However, this outbreak mainly originated at antigovernment rallies related to religious groups that avoided contact tracing [13] and thus did not reflect characteristics of typical community transmission in South Korea. Fourth, we did not include the genetic sequencing data of the cases, as genetic sequencing was not conducted in most cases. Therefore, we could not rule out a pseudo-outbreak of COVID-19 in which the PCR results were false positive [14] or in which infection was acquired from outside the cluster.

In conclusion, our study suggests that there were temporal changes in the possible risk of SSEs in South Korea. Ongoing monitoring of the risk of SSEs would help to assess the impact

of public health measures implemented in the community and to provide data for policy decision-making.

### Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

### Notes

**Disclaimer.** The funding bodies had no role in the study design, data collection and analysis, preparation of the manuscript, or decision to publish.

**Financial support.** This work was supported by the Basic Science Research Program through the National Research Foundation of Korea, funded by the Korean Ministry of Education (NRF-2020R111A3066471).

**Potential conflicts of interest.** All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

### References

1. Sun K, Wang W, Gao L, et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* **2021**; 371:eabe2424.
2. Ali ST, Wang L, Lau EHY, et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. *Science* **2020**; 369:1106–9.
3. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* **2005**; 438:355–9.
4. Endo A, Abbott S, Kucharski AJ, Funk S; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res* **2020**; 5:67.
5. Ryu S, Ali ST, Noh E, et al. Transmission dynamics and control of two epidemic waves of SARS-CoV-2 in South Korea. *BMC Infect Dis* **2021**; 21:485.
6. Kucharski AJ, Althaus CL. The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission. *Euro Surveill* **2015**; 20:14–8.
7. Althouse BM, Wenger EA, Miller JC, et al. Superspreading events in the transmission dynamics of SARS-CoV-2: opportunities for interventions and control. *PLoS Biol* **2020**; 18:e3000897.
8. Adam DC, Wu P, Wong JY, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat Med* **2020**; 26:1714–9.
9. Yonhap New Agency. S. Korea to test QR codes at nightclubs, eateries, cinemas to contain virus. **2020**. <https://en.yna.co.kr/view/AEN20200601003000315>. Accessed 10 February 2021.
10. The Korea Times. Ban on religious gatherings. Seoul: The Korea Times; **2020**.
11. Korean Ministry of Health and Welfare. Updates on COVID-19 in Republic of Korea. **2020**. [http://ncov.mohw.go.kr/tcmBoardView.do?brdId=&brdGubun=&ataGubun=&ncvContSeq=355454&contSeq=355454&board\\_id=&gubun=ALL](http://ncov.mohw.go.kr/tcmBoardView.do?brdId=&brdGubun=&ataGubun=&ncvContSeq=355454&contSeq=355454&board_id=&gubun=ALL). Accessed 10 February 2021.
12. Kim HJ, Hwang HS, Choi YH, et al. The delay in confirming COVID-19 cases linked to a religious group in Korea. *J Prev Med Public Health* **2020**; 53:164–7.
13. Lee S-E. Seoul records first three-digit figure of daily Covid-19 cases. Seoul, Korea: Joongang Ilbo; **2020**.
14. Mandal S, Tatti KM, Woods-Stout D, et al. Pertussis pseudo-outbreak linked to specimens contaminated by *Bordetella pertussis* DNA from clinic surfaces. *Pediatrics* **2012**; 129:e424–30.