# Technical Appendix: Temporal changes of the risk of superspreading events of coronavirus diseases 2019

# Authors:

Jun-Sik Lim1,2, Eunbi Noh1,3, Eunha Shim4, Sukhyun Ryu3

**Affiliations:**1. Graduate School of Public Health, Seoul National University, Seoul, 08826, Republic of Korea

2. College of Veterinary Medicine and Institute of Veterinary Science, Kangwon National University, Chuncheon, Korea

3. Department of Preventive Medicine, Konyang University College of Medicine, Daejeon, 35365, Republic of Korea

4. Department of Mathematics, Soongsil University, Seoul, 06978, Republic of Korea

# Author for correspondence: Prof. Sukhyun Ryu

Department of Preventive Medicine, Konyang University College of Medicine, R707, Myungok-Euihak Gwan, 158, Gwanjeodong-ro, Seogu, Daejeon, South Korea.
Telephone: +82 42 600 8673; Email: gentryu@onehealth.or.kr

**Word count (abstract):** 42/50; **Word count (main text):** 1054/1500

Running title: Change in the risk of superspreading of COVID-19

# Method

Based on the regression model, the probability that an index case ( for period () generate the number of secondary case () is given by

where and is the estimated dispersion parameter and reproduction number for period .

This model incorporates the covariates into the and as log-linear model like follows:

where and are intercepts, and are vectors of coefficients of period variable, and are vectors of period variable for and , respectively.

To understand the transmission dynamics of COVID-19 on different two epidemic period, two analyses were conducted: (1) analysis for only period variable, (2) analysis for period variable adjusted with the sex and age of infector. In analysis for only period variable, only period variable was incorporated in both and parts of negative binomial regression; In analysis for period variable adjusted with the sex and age of infector, sex and age of infector as well as period variable was incorporated in both and parts of negative binomial regression.

Given the estimated reproduction number and dispersion parameter for each period, the proportion of infected person responsible for 80% of secondary cases was calculated using equations from the previous studies [1-3]. The proportion is given by

where satisfies

Moreover, the risk of being infector and non-infector

Furthermore, with the threshold for SSE as 6 secondary cases defined in the previous study [1], the proportion of SSE was estimated for each period with the equations like below :

Finally, using the branching process, the expected probability that one index case of SARS-CoV-2 infection results in a cluster of size were estimated with the equations from previous studies [2, 3] given by

where the probability of cluster of size or greater could be estimated like below

Following the condition that seed cases were introduced into the totally susceptible populations, the estimated probability that at least one cluster of size or greater occurs is

Note that is the same with with the condition that one index case was introduced ().

We used a Bayesian Markov Chain Monte Carlo (MCMC) sampling. Four chains of 50,000 iterations were obtained with 20,000 burn-in. Prior distribution for the coefficients in each regression were set as normal distribution with mean of 0 and variance of 100,000. Convergence was checked by the trace plot visually and Gelman-Rubin-Brooks diagnostic [4]. The posterior distribution of the estimates is demonstrated with the median and the 95% credible intervals (95% CrI).

# References

1. Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nat Med. 2020;26(11):1714-9.

2. Endo A, Centre for the Mathematical Modelling of Infectious Diseases C-WG, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Res. 2020;5:67.

3. Kucharski AJ, Althaus CL. The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission. Euro Surveill. 2015;20(25):14-8.

4. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis: CRC press; 2013.