Time between infections versus time between symptom onset in COVID-19: implications for estimating the reproduction number

Jesse Knight[1,2] and Sharmistha Mishra[1,2,3,4]

[1]Institute of Medical Science, University of Toronto
[2]MAP Centre for Urban Health Solutions, Unity Health Toronto
[3]Division of Infectious Disease, Department of Medicine, University of Toronto
[4]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto

CAIMS 2021 Mini-symposium: *Mathematical modeling of COVID-19 transmission and mitigation strategies: efforts to end the pandemic*

Background: The effective reproduction number $R_e(t)$ is the average number of new infections directly generated from each existing infection under the conditions at time t. $R_e(t)$ can be calculated from the number of daily infections and the time between subsequent infections—the infection-infection distribution, $G(t)$. The infection-infection distribution $G(t)$ is often approximated by the symptom-symptom distribution $S(t)$, because the time of infection can be difficult to determine. However, if the time between infection and symptom onset—the infection-symptom distribution $H(t)$—varies substantially, then the infectee can develop symptoms before the infector and $S(t)$ can be negative, such as in the case of COVID-19. In this case, it may be improper to approximate $G(t)$ with $S(t)$.

Methods: Given parametric equations for the symptom-symptom distribution $S(t)$ and the infection-symptom distribution $H(t)$, we develop a method to recover the infection-infection distribution $G(t)$ using approximate deconvolution. We then compare estimates of $R_e(t)$ for the Greater Toronto Area using $G(t)$ to those using $S(t)$; two definitions of $S(t)$ are considered, which do and do not allow negative values, respectively.

Results: We estimated the time between COVID-19 infections $G(t)$ to be Gamma-distributed with mean 4.08 and standard deviation 3.19 days. The negative-permitting distribution $S(t)$ had equal mean but larger variance than $G(t)$, resulting in underestimation of $R_e(t)$ relative to $G(t)$, whereas the non-negative $S(t)$ had similar variance but larger mean, resulting in overestimation of $R_e(t)$.

Discussion: Approximation of the infection-infection distribution $G(t)$with the symptom-symptom distribution $S(t)$ may result in biased estimates of the effective reproduction number $R(t)$. The infection-infection distribution $G(t)$ can also be understood as the distribution of infectiousness; thus accurately distinguishing $G(t)$ from $S(t)$ may also have implications for isolation interventions. Future work should explore possible correlation between $S(t)$, $H(t)$, and $G(t)$ and estimation of confidence intervals for distribution parameters.

Open Science: The daily GTA case data used are not public. All other analysis code and data are available at: https://github.com/mishra-lab/covid-r