Biases in early-outbreak estimates of epidemiological delay distributions: applications to COVID-19 outbreak Sang Woo Park

Abstract

1 Introduction

Since the emergence of the novel coronavirus disease (COVID-19), a significant amount of research has focused on estimating relevant epidemiological parameters, particularly those describing time delays between key epidemiological events (Backer et al., 2020; Du et al., 2020; Ganyani et al., 2020; Lauer et al., 2020; Li et al., 2020; Linton et al., 2020; Nishiura et al., 2020; Tian et al., 2020; Zhao et al., 2020). These events can be compared within an infected individual or between transmission pairs. Estimates of within-individual delays, such as the incubation period, allow us to determine the appropriate duration of quarantine for suspected cases. On the other hand, estimates of between-individual delays, such as serial (i.e., the time between symptom onset of transmission pairs) and generation (i.e., time between infection of transmission pairs) intervals, allow us to determine epidemic potential and thus the required amount of intervention. Therefore, biases in the estimates of the delay distributions will necessarily lead to biases in the conclusion that depends on the distribution.

Measuring a time delay between two epidemiological events typically requires observing both events. A delay between two events cannot be measured if either event has not occurred or has not been observed yet. Here, we show that this dependency can systematically bias the estimate of a delay distribution if it is not explicitly taken into account; this bias applies to *all* epidemiological delay distributions. We compare two nonparametric approaches and a likelihood-based approach for correcting the bias. We then apply these methods to evaluate the amount of potential bias present in the early-outbreak estimate of the mean incubation period.

References

Backer, J. A., D. Klinkenberg, and J. Wallinga (2020). Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance 25*(5).

Du, Z., X. Xu, Y. Wu, L. Wang, B. J. Cowling, and L. A. Meyers (2020). Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. *Emerging Infectious Diseases 26*(6).

Ganyani, T., C. Kremer, D. Chen, A. Torneri, C. Faes, J. Wallinga, and N. Hens (2020). Estimating the generation interval for COVID-19 based on symptom onset data. *medRxiv*.

Lauer, S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*.

Li, Q., X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*.

Linton, N. M., T. Kobayashi, Y. Yang, K. Hayashi, A. R. Akhmetzhanov, S.-m. Jung, B. Yuan, R. Kinoshita, and H. Nishiura (2020). Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: A statistical analysis of publicly available case data. *Journal of Clinical Medicine 9*(2), 538.

Nishiura, H., N. M. Linton, and A. R. Akhmetzhanov (2020). Serial interval of novel coronavirus (COVID-19) infections. *International Journal of Infectious Diseases*.

Tian, S., N. Hu, J. Lou, K. Chen, X. Kang, Z. Xiang, H. Chen, D. Wang, N. Liu, D. Liu, et al. (2020). Characteristics of COVID-19 infection in Beijing. *Journal of Infection*.

Zhao, S., D. Gao, Z. Zhuang, M. Chong, Y. Cai, J. Ran, P. Cao, K. Wang, Y. Lou, W. Wang, et al. (2020). 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*.