Assessing uncertainties associated with early estimates of the basic reproduction number during the novel coronavirus (2019-nCoV) outbreak

Sang Woo Park[1] Jonathan Dushoff[2,3]

1 Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, USA
2 Department of Biology, McMaster University, Hamilton, Ontario, Canada
3 Michael G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada
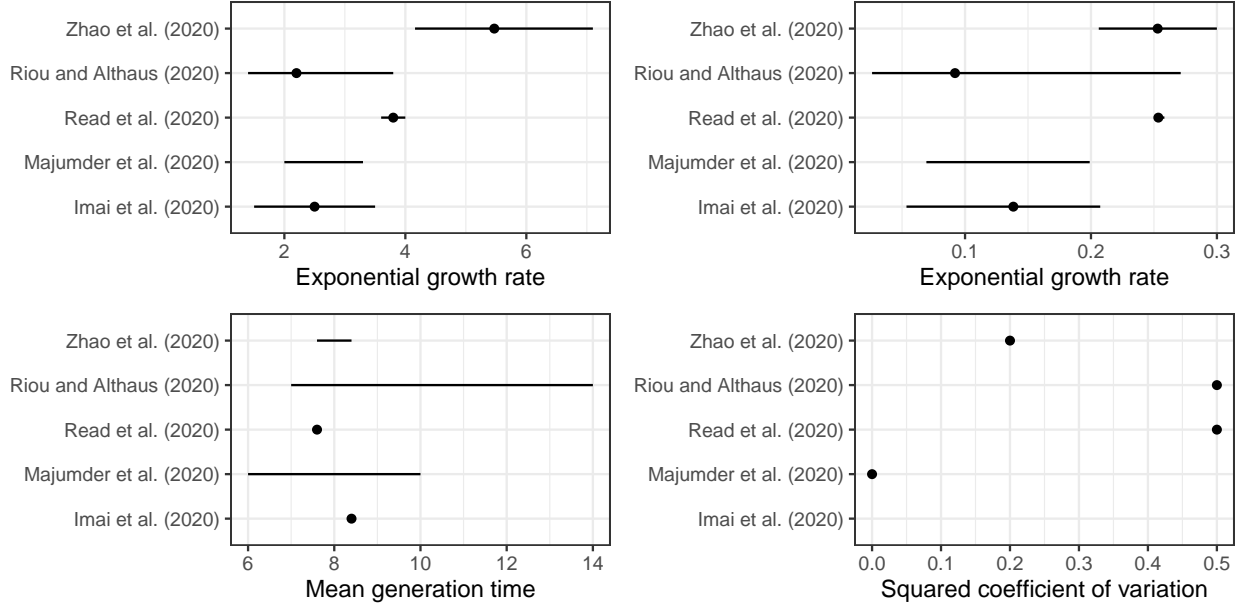
Abstract

Figure 1:   Early estimates of $\mathcal{R}_0$ and associated assumptions about r and $g(\tau)$.

As the novel coronavirus (2019-nCoV) continues to spread, researchers are rushing to publish their estimates of the basic reproductive number $\mathcal{R}_0$. The basic reproductive number (i.e., the average number of secondary cases generated by a primary case in a fully susceptible population) is of particular interest because it allows prediction about the final size of an epidemic. While their efforts are valuable, their analyses rely on several assumptions that could immediately affect their estimates of $\mathcal{R}_0$ and the associated uncertainties in their estimates. Here, we illustrate the degree to which their assumptions can affect their estimates and stress few principles that needs to be taken into consideration.

Early estimates of $\mathcal{R}_0$

Early in an outbreak, $\mathcal{R}_0$ cannot be estimated directly, whereas the exponential growth rate r can be estimated reliably. Given estimates of the exponential growth rate r and the distribution $g(\tau)$ of generation intervals (i.e., the time between when a person become infected and that person infects another person), the basic reproduction number can be estimated via the Euler-Lotka equation:

$$1/\mathcal{R}_0 = \int \exp(-r\tau)g(\tau)d\tau. \tag{1}$$

Therefore, early estimates of $\mathcal{R}_0$ must necessarily depend on the assumptions about r and $g(\tau)$.

    We use the gamma approximation framework to demonstrate how assumptions about r and $g(\tau)$ affects $\mathcal{R}_0$. Assuming that generation intervals follow a gamma distribution with

the mean $\bar{G}$ and the squared coefficient of variation κ, we have

$$\mathcal{R}_0 = \left(1 + \kappa r \bar{G}\right)^{1/\kappa}. \tag{2}$$

This equation demonstrates that generation-interval distributions with a larger mean (higher $\bar{G}$) or less variability (lower κ) will give a higher estimate of $\mathcal{R}_0$. Likewise, strong assumptions about $\bar{G}$ and κ will give estimates of $\mathcal{R}_0$ with narrow confidence intervals.

Figure 1 clearly demonstrates the lack of uncertainties in the underlying parameters of the current $\mathcal{R}_0$ estimates. In particular, no studies properly propagate uncertainties associated with the amount of variability in generation intervals, which can have large effects on the estimates of $\mathcal{R}_0$ For example, when the relative mean generation interval is long ($r\bar{G} \approx 2$ based on the estimates by Zhao et al. (2020)), varying κ from 0 to 1 reduces the estimate of $\mathcal{R}_0$ by more than a two fold.

To assess the amount of uncertainties in the estimate of $\mathcal{R}_0$, we come up with distributions that span over the assumptions made by other people:

$$
\begin{aligned}
r &\sim \text{Gamma}(\alpha = 5, \beta = 5/0.15) \\
\bar{G} &\sim \text{Gamma}(\alpha = 10, \beta = 10/8.5) \\
\kappa &\sim \text{Gamma}(\alpha = 10, \beta = 10/0.35)
\end{aligned}
\tag{3}
$$

These give median estimate of $\mathcal{R}_0$ of 2.37 with 95% qunatile of 1.26–9.01.

Study	$\mathcal{R}_0$	r (days^{-1})	$\bar{G}$ (days)	κ
Imai et al. (2020)	2.5 (1.5–3.5)	0.14 (0.05–0.21)	8.4	unspecified*
Majumder and Mandl (2020)	2.0–3.3	0.07–0.20	6–10	0
Read et al. (2020)	3.8 (3.6–4.0)	0.25 (0.25–0.26)[†]	7.6	0.5
Riou and Althaus (2020)	2.2 (1.4–3.8)	0.09 (0.03–0.27)	7–14	0.5
Zhao et al. (2020)	5.47 (4.16–7.1)[‡]	0.25 (0.21–0.30)	7.6–8.4	0.2

Table 1: **Parameter estimates and assumptions** Exponential growth rate r is calculated by using the following formula: $r = (\mathcal{R}_0^\kappa - 1)/(\kappa \bar{G})$. *We assume $\kappa = 0.5$ to calculate r. [†]Exponential growth rate r is calculated by using the exact formula for the SEIR model (Ma et al., 2014). [‡]Assuming no increase in the reporting rate.

References

Imai, N., A. Cori, I. Dorigatti, M. Baguelin, C. A. Donelly, S. Riley, and N. M. Ferguson (2020). Report 3: Transmissibility of 2019-nCoV. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-2019-nCoV-transmissibility.pdf.

Ma, J., J. Dushoff, B. M. Bolker, and D. J. Earn (2014). Estimating initial epidemic growth rates. *Bulletin of mathematical biology 76*(1), 245–260.

Majumder, M. and K. D. Mandl (2020). Early transmissibility assessment of a novel coronavirus in Wuhan, China. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3524675.

Read, J. M., J. R. Bridgen, D. A. Cummings, A. Ho, and C. P. Jewell (2020). Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. https://www.medrxiv.org/content/10.1101/2020.01.23.20018549v1.

Riou, J. and C. L. Althaus (2020). Pattern of early human-to-human transmission of wuhan 2019-nCoV. https://github.com/jriou/wcov/tree/5b921454393f130f33a766b42ae2474749719cb1.

Zhao, S., J. Ran, S. S. Musa, G. Yang, Y. Lou, D. Gao, L. Yang, and D. He (2020). 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. https://www.biorxiv.org/content/10.1101/2020.01.23.916395v1.