The generation-interval distribution of asymptomatic transmission shapes the epidemic potential and the accuracy of $\mathcal{R}_0$ assessments in COVID-19 outbreak

Sang Woo Park,[1] Daniel M. Cornforth,[2] Jonathan Dushoff,[3,4,5,*] and Joshua S. Weitz[2,6,†]

[1] *Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA*

[2] *School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA*

[3] *Department of Biology, McMaster University, Hamilton, Ontario, Canada*

[4] *Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada*

[5] *M. G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada*

[6] *School of Physics, Georgia Institute of Technology, Atlanta, GA, USA*

(Dated: March 7, 2020)

We assess the impact of asymptomatic cases on epidemic potential of COVID-19 (and COVID-19 like pathogens) – as measured both by the basic reproduction number (i.e., the expected number of secondary cases generated by an average primary case in a fully susceptible population) and the fraction of new secondary cases attributable to asymptomatic individuals. As we show, the impact of asymptomatic cases depends on their generation intervals (i.e., time between when an individual is infected and when that individual infects another person). If the generation intervals of asymptomatic cases differs from that of symptomatic cases, then estimates of the basic reproduction number which do not explicitly account for asymptomatic cases may be systematically biased. Specifically, if asymptomatic cases have a shorter generation interval, $\mathcal{R}_0$ will be over-estimated, and if they have a longer generation interval, $\mathcal{R}_0$ will be under-estimated. The analysis provides a rationale for assessing the duration of asymptomatic cases of COVID-19 in addition to their prevalence in the population.

*Electronic address: dushoff@mcmaster.ca; URL: https://mac-theobio.github.io/dushoff.html

I. INTRODUCTION

In an epidemic, symptomatic cases are the predominant focus of treatment and usually represent the bulk of reported cases. However, infected individuals who are asymptomatic yet infectious can be a critical factor in the spread of some pathogens [1]. Asymptomatic individuals are hard to trace, unlikely to self-isolate, and are likely to retain normal social and travel patterns. The COVID-19 outbreak has raised significant questions regarding the role of asymptomatic cases [2]. A key focus has been on estimating asymptomatic prevalence (see [3]), e.g., what fraction of individuals have asymptomatic infections relative to those with symptomatic infections?

There is significant ongoing interest in asymptomatic infections in COVID-19 for two major reasons. First, the proportion of infections that are asymptomatic is critical to attempts to estimate the likely burden of severe outcomes (including mortality) when the virus spreads through a population. Second, understanding the possible role of *transmission* by asymptomatic individuals is crucial to planning surveillance and control efforts.

Here, we focus on a third effect. If asymptomatic cases are important for transmission, they also have the potential to affect estimates of key parameters of disease spread such as the basic reproduction number $\mathcal{R}_0$ (i.e., the expected number of secondary cases generated by an average primary case in a fully susceptible population [4]). Here, we investigate the relationship between individual-level features of asymptomatic cases (e.g., the probability of initiating an asymptomatic case, asymptomatic case duration, and the transmission rates from asymptomatic individuals), which are represented by a generation-interval distribution, to dynamics at the population scale. That is, rather than focusing on the prevalence of asymptomatic cases, we are interested in understanding the relevance of asymptomatic infections to secondary case production.

II. METHODS

We model viral spread using a renewal-equation framework [5], which allows us to model the current incidence of infected individuals (i.e., the the rate at which new infections occur

[†]Electronic address: jsweitz@gatech.edu; URL: http://ecotheory.biology.gatech.edu

in the population) as a function of previous incidence and how infectiousness of an infected individual varies over the course of their infection. We divide incidence i into two categories – i_a and i_s – corresponding to incidence of asymptomatic and symptomatic cases, respectively. Newly infected individuals that are either asymptomatically or symptomatically infected can transmit the disease to others, but they differ in their intrinsic reproduction numbers, $\mathcal{R}_a$ and $\mathcal{R}_s$, respectively, as well as intrinsic generation-interval distributions, $g_a(\tau)$ and $g_s(\tau)$. Generation intervals, which are defined as the time between when an individual is infected and when that individual infects another person [6], depend on the natural history of infection: individuals with subclinical infections may have fast clearance and short generation intervals, or slow viral reproduction and long generation intervals (cf. [7]). The shape of the generation-interval distribution characterizes the relationship between the epidemic growth rate r and the reproduction number [8].

Neglecting births and loss of immunity on the time scale of the outbreak, the dynamics of susceptibles and incidence are:

$$\dot{S} = -i(t) \tag{1}$$

$$i(t) = \mathcal{R}_a S(t) \int_0^\infty i_a(t-\tau) g_a(\tau) \mathrm{d}\tau + \mathcal{R}_s S(t) \int_0^\infty i_s(t-\tau) g_s(\tau) \mathrm{d}\tau. \tag{2}$$

The basic reproduction number of this system is:

$$\mathcal{R}_0 = p\mathcal{R}_a + (1-p)\mathcal{R}_s, \tag{3}$$

where p is the proportion of *incident cases* that are asymptomatic: $i_a(t) = pi(t)$. We define the "intrinsic" proportion of asymptomatic transmission as the relative contribution of asymptomatic cases to the basic reproduction number:

$$z = p\mathcal{R}_a/\mathcal{R}_0. \tag{4}$$

Note that the intrinsic proportion of symptomatic transmission satisfies

$$1 - z = (1-p)\mathcal{R}_s/\mathcal{R}_0. \tag{5}$$

The intrinsic generation-interval distribution [9] depends on the intrinsic proportion of asymptomatic vs. symptomatic transmission and their corresponding intrinsic generation-interval distributions:

$$g(\tau) = z g_a(\tau) + (1-z) g_s(\tau), \tag{6}$$

which allows us to express incidence of infection as a renewal process that depends on previous incidence:

$$i(t) = \mathcal{R}_0 S(t) \int_0^\infty i(t-\tau) g(\tau) \mathrm{d}\tau. \tag{7}$$

Yet, this information is not sufficient to disentangle the role of asymptomatic cases, i.e., what fraction of secondary cases can be ascribed to *realized* transmission from asymptomatic cases vs. symptomatic cases?

We note that, in practice, counted cases may largely be of the symptomatic type. Hence, we ask: is the expected ratio of incidence caused by asymptomatic cases vs. symptomatic cases during the outbreak phase the same or different from $z/(1-z)$? The ratio of secondary case production caused by asymptomatic vs. symptomatic individuals over the course of an ongoing epidemic is

$$\frac{q}{1-q} = \frac{\mathcal{R}_a S(t) \int_0^\infty i_a(t-\tau) g_a(\tau) \mathrm{d}\tau}{\mathcal{R}_s S(t) \int_0^\infty i_s(t-\tau) g_s(\tau) \mathrm{d}\tau}, \tag{8}$$

where q is the fraction of new secondary cases caused by asymptomatic individuals or 'relevance' of asymptomatic cases. Assuming that the incidence grows exponentially during the outbreak phase (i.e., $i(t) = i(0)\exp(rt)$) and substituting $i_a(t) = pi(t)$ and $i_s(t) = (1-p)i(t)$, we have

$$\frac{q}{1-q} = \left(\frac{z}{1-z}\right) \left[\frac{\int_0^\infty \exp(-r\tau) g_a(\tau) \mathrm{d}\tau}{\int_0^\infty \exp(-r\tau) g_s(\tau) \mathrm{d}\tau}\right]. \tag{9}$$

The value of q is an unknown feature of outbreaks at the population scale, distinct from p the probability of becoming asymptomatic upon infection or z the intrinsic proportion of asymptomatic transmission.

In order to understand how the shape of asymptomatic and symptomatic generation-interval distributions affects q and $\mathcal{R}_0$, we summarize the distributions in terms of their means and dispersions. Assuming that the intrinsic generation-interval distributions for asymptomatic and symptomatic cases follow gamma distributions with different means, $\bar{G}_a$ and $\bar{G}_s$, and dispersions (represented by squared coefficients of variations), κ_a and κ_s [10], we have

$$\frac{q}{1-q} = \left(\frac{z}{1-z}\right) \left[\frac{(1+\kappa_s r \bar{G}_s)^{1/\kappa_s}}{(1+\kappa_a r \bar{G}_a)^{1/\kappa_a}}\right]. \tag{10}$$

As is apparent from Eq. (9), relevance q increases with z, i.e., with both p and $\mathcal{R}_a/\mathcal{R}_s$. The consequences for outbreak potential further depends on the relative lengths of generation intervals as well as the relative amount of dispersions in them. In Supplementary Materials, we use a compartmental ordinary differential equation model to show that the duration of

infectious period of asymptomatic individuals affects both q and $\mathcal{R}_0$, which is a special case of Eq. (9); we also show that prevalence of asymptomatic/symptomatic cases differs from p and $1-p$.

III. RESULTS

We evaluate the impacts of intrinsic proportion of asymptomatic transmission z on the realized relevance of asymptomatic infection q and the basic reproduction number $\mathcal{R}_0$, given variation in relative mean generation interval $\bar{G}_a/\bar{G}_s$ of asymptomatic cases, using parameter values that are consistent with earlier COVID-19 models [11]: $1/r = 7$ days, $\bar{G}_s = 8$ days, and $\kappa_s = \kappa_a = 0.5$. We vary $\bar{G}_a/\bar{G}_s$ between 0.5 and 2, and infer values of q and $\mathcal{R}_0$ compatible with an estimated growth rate r from case data given the generation interval distributions of asymptomatic and symptomatic cases. Across the ranges of parameters we

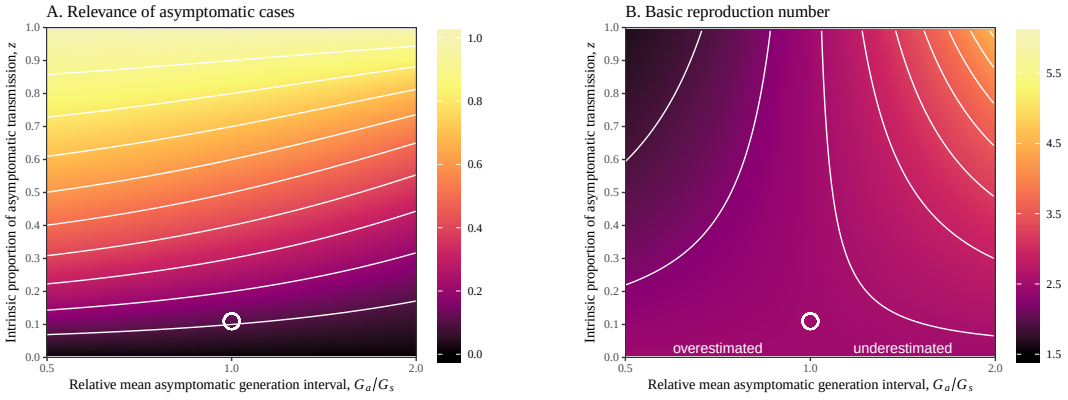

FIG. 1: Effects of intrinsic proportion of asymptomatic transmission on the relevance of asymptomatic transmission and basic reproduction number, given variation in the mean generation interval of asymptomatic cases. (A) Long/short generation intervals of asymptomatic cases decrease/increase the relevance of asymptomatic cases, q. (B) Long/short generation intervals of asymptomatic cases increase/decrease the basic reproduction number $\mathcal{R}_0$. We vary G_a/G_s while assuming $1/r = 7$ days, $\bar{G}_s = 8$ days, and $\kappa_s = \kappa_a = 0.5$. The white lines represent the contours. The white circles represent the baseline value that does not explicitly account for asymptomatic transmission; depending on whether G_a is smaller or bigger than G_s, the baseline value over- or under-estimates the true $\mathcal{R}_0$.

explore, the intrinsic proportion of asymptomatic transmission z is similar to the realized proportion q (Figure 1A). As the relative mean generation interval of asymptomatic cases, $\bar{G}_a/\bar{G}_s$, increases, q decreases because symptomatic cases are more likely to have short generation intervals (i.e., fast transmission events), which drive the spread during the growth phase (Figure 1A).

Figure 1B compares the effect of relative mean generation interval of asymptomatic cases, $\bar{G}_a/\bar{G}_s$, on the basic reproduction number $\mathcal{R}_0$. When $\bar{G}_s < \bar{G}_a$, then $\mathcal{R}_0$ increases with z. This increase arises because asymptomatic infections that have long generation intervals increase the mean of overall generation intervals (see [10]). Conversely, when $\bar{G}_s > \bar{G}_a$ then the strength $\mathcal{R}_0$ decreases with z because short generation intervals of asymptomatic cases reduce the overall mean generation interval. The qualitative effects of z and $\bar{G}_a/\bar{G}_s$ on q and $\mathcal{R}_0$ remain robust when we assume narrower ($\kappa_s = \kappa_a = 0.3$; Figure S1) or wider ($\kappa_s = \kappa_a = 0.8$; Figure S2) generation intervals.

As an example, we assume that 18% of the newly infected cases are asymptomatic [3] and that the asymptomatic cases are as half as infectious as the symptomatic ones (i.e., $\mathcal{R}_a = 0.5\mathcal{R}_s$). The intrinsic proportion of asymptomatic transmission z is 0.11, whereas asymptomatic cases can account for anywhere between 6% and 16% of infection during the exponential growth phase across the range of mean asymptomatic generation interval we explore (Figure 1A). Likewise, the basic reproduction number ranges between 2.3 and 2.6 (Figure 1B). Therefore, when $\mathcal{R}_0$ is estimated without explicitly accounting for asymptomatic spread ($\mathcal{R}_0 = 2.5$; red circle in Figure 1B), it can be over- or under- estimated depending on the relative duration of infection between symptomatic and asymptomatic individuals.

Relative generation-interval dispersion of asymptomatic cases κ_a/κ_s also have qualitatively similar effects on q and $\mathcal{R}_0$. When the asymptomatic cases have a wider/narrower generation-interval distribution, higher intrinsic proportion of asymptomatic transmission z increases/decreases the relevance of asymptomatic transmission q but decreases/increases the basic reproduction number (Figure S3). Since a wider generation-interval distribution has a higher proportion of early transmission than a narrow one, increasing the generation-interval dispersion has qualitatively similar effects on q and $\mathcal{R}_0$ as decreasing the mean generation interval.

IV. DISCUSSION

Much is still unknown about course of asymptomatic infection compared to symptomatic infection in COVID-19. Our current work highlights the need to characterize the duration of asymptomatic infections, relative to that of symptomatic ones, and their consequences not only for contact tracing but for precise estimation of the basic reproduction number of the ongoing COVID-19 outbreak [11]. The present findings are also consistent with a recent generalization linking speed, strength, and generation intervals, such that for a given observed speed increases in the mean generation interval imply larger reproduction number [10].

If mild, asymptomatic infections are more persistent than severe, symptomatic ones, the mean generation interval for COVID-19 could be longer than estimated from symptomatic cases alone - possibly increasing $\mathcal{R}_0$ (Figure 1B). However, if asymptomatic cases tend to have shorter infection, then current estimates of $\mathcal{R}_0$ may be over-estimates of the underlying strength (Figure 1B), although asymptomatic cases may be driving a larger fraction of secondary cases than we would expect without accounting for their differences (Figure 1A). Therefore, current estimates of $\mathcal{R}_0$ that do not account for asymptomatic cases may be systematically biased.

Finally, it is important to note we focused here on the exponential phase, but that the asymptomatic relevance q is time-dependent, and will vary given the population state structure of infected individuals. Such variation in the ratio of asymptomatic to symptomatic cases will also impact estimates of case fatality rates insofar as many asymptomatic cases are not counted, particularly early in an outbreak. Future work might also consider the ways in which asymptomatic individuals modulate resurgent epidemics in a networked metapopulation [12]. Characterizing the role of asymptomatic individuals in driving the persistence of the epidemic will be critical for assessing the post-pandemic outcome [13].

Acknowledgments: The authors thank John Glasser for comments and discussion on the manuscript, particularly on short notice. This work was supported, in part, by support from the Army Research Office to JSW (W911NF1910384).

Data Availabililty: All code is available at https://github.com/mac-theobio/coronavirus_asymptomatic.

Supplementary Materials

A compartmental model for asymptomatic/symptomatic cases

Consider an SEIR model variant in which an infected individual can be either asymptomatic, I_a, or symptomatic, I_s. We note that I_a and I_s represent prevalence (i.e., the total number of currently infectious individuals) of asymptomatic and symptomatic individuals; these quantities are different from i_s and i_s that we present in the main text, which represent incidence (i.e., the rate at which new cases are generated) of asymptomatic and symptomatic individuals. While both cases can recover, we assume that only symptomatic cases can lead to fatalities, denoted by the D category. In total, the dynamics of susceptibles, exposed, infectious, recovered, and dead are:

$$\dot{S} = -\beta_a S I_a - \beta_s S I_s \tag{S1}$$

$$\dot{E} = \beta_a S I_a + \beta_s S I_s - \gamma_e E \tag{S2}$$

$$\dot{I}_a = p\gamma_e E - \gamma_a I_a \tag{S3}$$

$$\dot{I}_s = (1-p)\gamma_e E - \gamma_s I_s \tag{S4}$$

$$\dot{R} = \gamma_a I_a + (1-f)\gamma_s I_s \tag{S5}$$

$$\dot{D} = f\gamma_s I_s. \tag{S6}$$

Here, β_a and β_s denote transmission rates, γ_e denotes the transition from exposed to infectious, p is the fraction of asymptomatic cases that are generated for each exposed individual, $1-p$ is the fraction of symptomatic cases that are generated for each exposed individual, γ_a and γ_s denote recovery rates, and f denotes the case fatality ratio for symptomatic cases.

Given that the number of infected individuals increase exponentially at rate r initially, the equations for the infectious cases can be rewritten given the ansatz $E(t) = c_e e^{rt}$, $I_a(t) = c_a e^{rt}$, $I_s(t) = c_s e^{rt}$. Then, it follows that

$$rc_a = p\gamma_e c_e - \gamma_a c_a, \tag{S7}$$

$$rc_s = (1-p)\gamma_e c_e - \gamma_s c_s, \tag{S8}$$

which implies that

$$\frac{c_a}{c_s} = \frac{p}{1-p}\left[\frac{r+\gamma_s}{r+\gamma_a}\right]. \tag{S9}$$

The prevalence of asymptomatic and symptomatic individuals is different from p and $1-p$ because prevalence measures the individuals that are currently infectious and does not account for individuals that have already recovered. Finally, the ratio of secondary case production caused by asymptomatic vs. symptomatic individuals during the exponential phase should be

$$\frac{q}{1-q} = \left(\frac{\beta_a}{\beta_s}\right) \frac{p}{1-p} \left[\frac{r+\gamma_s}{r+\gamma_a}\right], \tag{S10}$$

where q is the fraction of new secondary cases caused by asymptomatic individuals.

The basic reproduction number of this system is:

$$\mathcal{R}_0 = p\mathcal{R}_a + (1-p)\mathcal{R}_s, \tag{S11}$$

where

$$\mathcal{R}_a = \frac{\beta_a}{\gamma_a}, \tag{S12}$$

$$\mathcal{R}_s = \frac{\beta_s}{\gamma_s}. \tag{S13}$$

The generation-interval distributions for asymptomatic and symptomatic individuals follow the same functional form as the corresponding generation-interval distribution for a single-type SEIR model since both asymptomatic and symptomatic individuals have exponentially distributed latent and infectious periods [14]:

$$g_a(\tau) = \frac{\gamma_e\gamma_a}{\gamma_e-\gamma_a}\left(\exp(-\gamma_a\tau) - \exp(-\gamma_e\tau)\right), \tag{S14}$$

$$g_s(\tau) = \frac{\gamma_e\gamma_s}{\gamma_e-\gamma_s}\left(\exp(-\gamma_s\tau) - \exp(-\gamma_e\tau)\right). \tag{S15}$$

It immediately follows that

$$\left(\frac{z}{1-z}\right)\left[\frac{\int_0^\infty \exp(-r\tau)g_a(\tau)\mathrm{d}\tau}{\int_0^\infty \exp(-r\tau)g_s(\tau)\mathrm{d}\tau}\right] = \left(\frac{\beta_a}{\beta_s}\right)\frac{p}{1-p}\left[\frac{r+\gamma_s}{r+\gamma_a}\right], \tag{S16}$$

where $z = p\mathcal{R}_a/\mathcal{R}_0$ and $1-z = (1-p)\mathcal{R}_s/\mathcal{R}_0$.

Supplementary figures

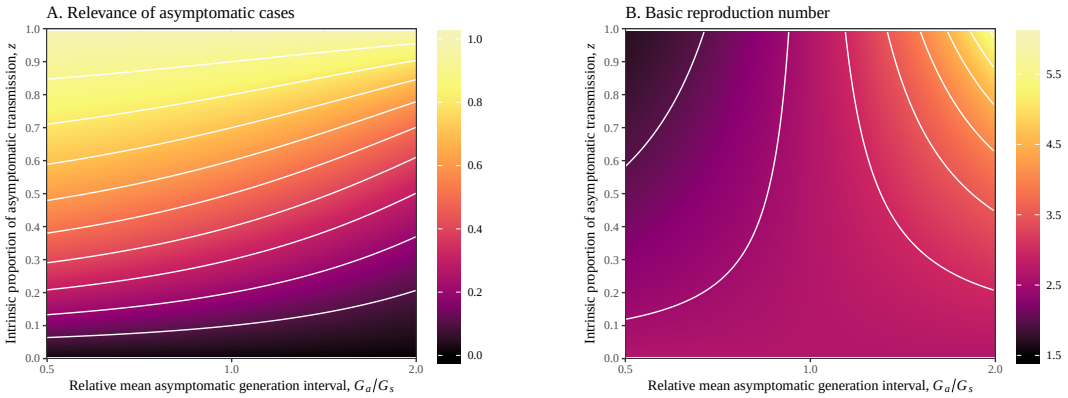

FIG. S1: Effects of intrinsic proportion of asymptomatic transmission on the relevance of asymptomatic transmission and basic reproduction number, given variation in the mean generation interval of asymptomatic cases when generation-interval distributions are narrow. We vary G_a/G_s while assuming $1/r = 7$ days, $\bar{G}_s = 8$ days, and $\kappa_s = \kappa_a = 0.3$. See Figure 1 in the main text for figure caption.

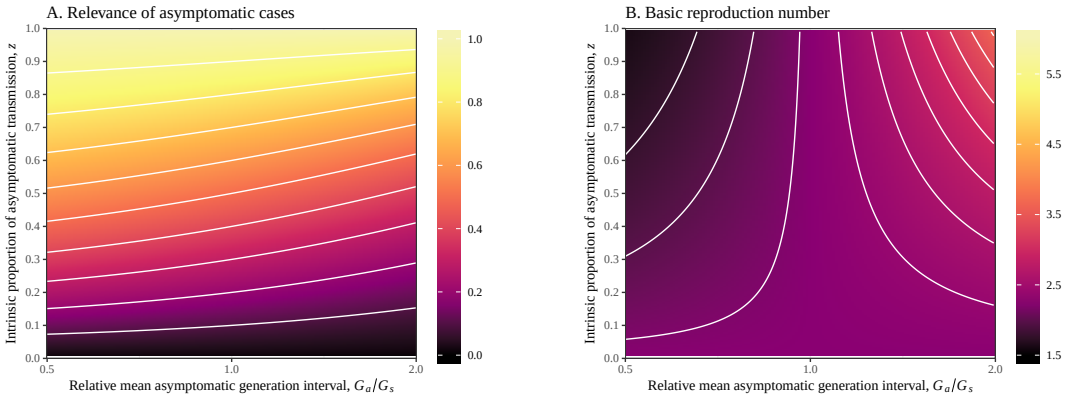

FIG. S2: Effects of intrinsic proportion of asymptomatic transmission on the relevance of asymptomatic transmission and basic reproduction number, given variation in the mean generation interval of asymptomatic cases when generation-interval distributions are wide. We vary G_a/G_s while assuming $1/r = 7$ days, $\bar{G}_s = 8$ days, and $\kappa_s = \kappa_a = 0.8$. See Figure 1 in the main text for figure caption.

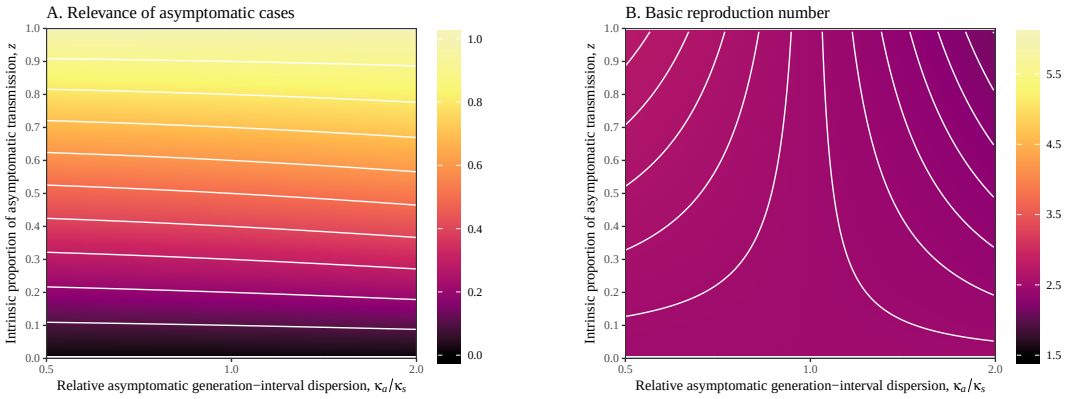

FIG. S3: Effects of intrinsic proportion of asymptomatic transmission on the relevance of asymptomatic transmission and basic reproduction number, given variation in the generation-interval dispersion of asymptomatic cases. (A) Wide/narrow generation intervals of asymptomatic cases increase/decrease the relevance of asymptomatic cases, q. (B) Wide/narrow generation intervals of asymptomatic cases decrease/increase the basic reproduction number $\mathcal{R}_0$. We vary κ_a/κ_s while assuming $1/r = 7$ days, $\bar{G}_s = \bar{G}_a = 8$ days, and $\kappa_s = 0.5$. The white lines represent the contours.

[1] Fraser, C., Riley, S., Anderson, R. M. & Ferguson, N. M. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences* **101**, 6146–6151 (2004).

[2] Fauci, A. S., Lane, H. C. & Redfield, R. R. Covid-19 – navigating the uncharted. *New England Journal of Medicine* (2020). URL `https://doi.org/10.1056/NEJMe2002387`.

[3] Mizumoto, K., Kagaya, K., Zarebski, A. & Chowell, G. Estimating the asymptomatic ratio of 2019 novel coronavirus onboard the Princess Cruises Ship, 2020. *medRxiv* (2020). URL `https://www.medrxiv.org/content/10.1101/2020.02.20.20025866v2.full`.

[4] Anderson, R. M. & May, R. M. *Infectious diseases of humans: dynamics and control* (Oxford university press, 1992).

[5] Heesterbeek, J. & Dietz, K. The concept of ∇_0 in epidemic theory. *Statistica Neerlandica* **50**, 89–110 (1996).

[6] Svensson, Å. A note on generation times in epidemic models. *Mathematical biosciences* **208**, 300–311 (2007).

[7] Roberts, M. & Heesterbeek, J. Model-consistent estimation of the basic reproduction number from the incidence of an emerging infection. *Journal of mathematical biology* **55**, 803 (2007).

[8] Wallinga, J. & Lipsitch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* **274**, 599–604 (2007).

[9] Champredon, D. & Dushoff, J. Intrinsic and realized generation intervals in infectious-disease transmission. *Proceedings of the Royal Society B: Biological Sciences* **282**, 20152026 (2015).

[10] Park, S. W., Champredon, D., Weitz, J. S. & Dushoff, J. A practical generation-interval-based approach to inferring the strength of epidemics from their speed. *Epidemics* **27**, 12–18 (2019).

[11] Park, S. W. *et al.* Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. *medRxiv* (2020). URL `https://www.medrxiv.org/content/10.1101/2020.01.30.20019877v4.full.pdf`.

[12] Watts, D. J., Muhamad, R., Medina, D. C. & Dodds, P. S. Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proceedings of the National Academy of Sciences* **102**,

11157–11162 (2005).

[13] Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the post-pandemic period. *medRxiv* (2020). URL `https://www.medrxiv.org/content/10.1101/2020.03.04.20031112v1`.

[14] Svensson, Å. The influence of assumptions on generation time distributions in epidemic models. *Mathematical biosciences* **270**, 81–89 (2015).