science.sciencemag.org/cgi/content/full/science.abb6936/DC1

Supplementary Materials for

Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing

Luca Ferretti, Chris Wymant, Michelle Kendall, Lele Zhao, Anel Nurtay, Lucie Abeler-Dörner, Michael Parker, David Bonsall, Christophe Fraser*

*Corresponding author. Email: christophe.fraser@bdi.ox.ac.uk

Published 31 March 2020 on *Science* First Release
DOI: 10.1126/science.abb6936

This PDF file includes:

Other supplementary material for this manuscript includes:

Data S1 (Excel file)

Supplementary Text

We note that this work was posted as pre-print on medrxiv (41).

Exponential growth rate

We took our point estimate of the exponential growth rate $r = 0.14$ per day ($T_2 = 5.0$ days) from (20). That value was derived from the exponential growth in the number of deaths per day in China from January 25th 2020 to Feb 4th 2020. To derive the CI used here we fit a linear model to the same data (42).

Inference of the distribution of generation times

The distribution of generation times was inferred by Maximum Composite Likelihood from serial intervals and exposure periods of 40 transmission pairs with known dates of onset of symptoms. Some of the 40 transmission pairs were taken from references (43), (44), (45), (12); some taken from a previous estimate of serial intervals (22); some we identified from other reports (see Supplementary Table). This set of transmission events spans different countries (see Supplementary Figure 5) and social relations (households, family members, colleagues, acquaintances, random contacts) and are included in a Supplementary file. For each transmission event $1 \rightarrow 2$, we include the dates of onset of symptoms $t_{s,1}$ and $t_{s,2}$, the intervals of exposure $[e_{1,L}, e_{1,R}]$ and

$[e_{2,L}, e_{2,R}]$, and the reporting date T_r. The intervals of exposure must satisfy the conditions $e_{1,R} \leq \min(t_{s,1}, e_{2,R})$, $e_{2,L} \geq e_{1,L}$ and $e_{2,R} \leq t_{s,2}$.

First, we discretize the incubation and generation time distributions

$$I(j) = \int_{j-0.5}^{j+0.5} i(t)\, dt\,, \qquad \Omega(j|\Theta_\omega) = \int_{j-0.5}^{j+0.5} \omega(t)\, dt \tag{1}$$

where the incubation time distribution $i(t)$ is taken from [Lauer et al 2020], and Θ_ω denotes the parameters of the generation time distribution $\omega(t)$. Then, for a given transmission event $1 \to 2$, we define the likelihood of observing the transmission and the subsequent time of onset of symptoms in the recipient as

$$\mathcal{L}_{\text{trans}}(\Theta_\omega) = P[1 \to 2, t_{s,2}|t_{s,1}, \Theta_\omega] \tag{2}$$

which can be obtained by summing over all possible infection times $t_{i,1}$ and $t_{i,2}$, resulting in the explicit form

$$\mathcal{L}_{\text{trans}}(\Theta_\omega) \propto \sum_{t_{i,1}=e_{1,L}}^{e_{1,R}} e^{-r(T_r - t_{i,1})} I(t_{s,1} - t_{i,1}) \sum_{t_{i,2}=e_{2,L}}^{e_{2,R}} \Omega(t_{i,2} - t_{i,1}|\Theta_\omega) I(t_{s,2} - t_{i,2}) \tag{3}$$

up to a multiplicative term independent of $\omega(t)$ which has no effect on likelihood maximisation. The growth rate r corresponds to a doubling time of 5 days for exponentially growing epidemics (China except Hong Kong, Italy, South Korea) while it is assumed to be 0 for countries where local transmission is limited.

For simplicity, we ignore correlations between transmission events. We define the composite likelihood across all 40 pairs (i.e. the approximate global likelihood that neglects correlations) as the product of the likelihoods $\mathcal{L}_{\text{trans}}(\Theta_\omega)$ of all pairs. We then test multiple functional forms for $\omega(t)$ (lognormal, gamma, Weibull), inferring their parameters $\hat{\Theta}_\omega$ by Maximum Composite Likelihood (MCL). Finally, we select the functional form based on the lowest AIC. Confidence intervals for the parameters of the distribution are based on likelihood ratios, while confidence intervals for the values of the curve are based on likelihood profiling.

This approach relies not only on the timing of transmissions and onset of symptoms, but it also assumes that transmission pairs themselves are chosen at random from the epidemic. This is not generally the case. To test the robustness of our approach, we relax this assumption by considering also the probability conditional on the occurrence of the transmission $1 \to 2$ at any time:

$$\mathcal{L}_{|\text{trans}}(\Theta_\omega) = P[t_{s,2}|1 \to 2, t_{s,1}, \Theta_\omega] = \frac{P[1 \to 2, t_{s,2}|t_{s,1}, \Theta_\omega]}{P[1 \to 2|t_{s,1}, \Theta_\omega]} = \frac{\mathcal{L}_{\text{trans}}(\Theta_\omega)}{\sum_{t'_{2,s}} P[1 \to 2, t'_{2,s}|t_{s,1}, \Theta_\omega]} \tag{4}$$

This approach does not extract any information from the fact that a transmission actually occurred. The MCL inference uses only the timing of exposure and onset of symptoms. The correspondence between the two approaches is therefore a good test of robustness.

Posterior probability of pre-symptomatic transmission

The probability that infection occurred before onset of symptoms for the infector was estimated for each transmission pair using a Bayesian approach based on the best fit for the generation time distribution (see Supplementary Information). This estimate assumes independence between generation time and incubation period, but takes into account the available information on

period of exposure and onset of symptoms for each case. In detail: for each transmission event, the likelihood $\mathcal{L}_{\text{trans}}(\Theta_\omega)$ described above can be decomposed as

$$\mathcal{L}_{\text{trans}} = \mathcal{L}_{\text{pre}} + \mathcal{L}_{\text{pre}} \tag{5}$$

where the pre-symptomatic term $\mathcal{L}_{\text{pre}}$ includes only the cases with $t_{i,2} < t_{s,1}$ and the symptomatic term $\mathcal{L}_{\text{sym}}$ includes all other cases. In the Main Text we assume a prior probability $p_{\text{pre}} = 1 - p_{\text{sym}} = 0.5$ for pre-symptomatic and symptomatic transmission, while in Supplementary we also include different prior beliefs ($p_{\text{pre}} = 0.25$ and $p_{\text{pre}} = 0.75$). Then, for each transmission event, we can obtain the posterior probability of pre-symptomatic transmission via a Bayesian approach:

$$P[\text{presymptomatic transmission}] = \frac{p_{\text{pre}}\mathcal{L}_{\text{pre}}(\hat{\Theta}_\omega)}{p_{\text{pre}}\mathcal{L}_{\text{pre}}(\hat{\Theta}_\omega) + p_{\text{sym}}\mathcal{L}_{\text{sym}}(\hat{\Theta}_\omega)} \tag{6}$$

where all likelihoods are evaluated at the parameters $\hat{\Theta}_\omega$ that maximise the composite likelihood.

The distribution of the fraction of transmissions that occurred before onset of symptoms can be estimated by assigning each event as pre-symptomatic or symptomatic at random according to its posterior probability. The empirical distribution of this quantity is obtained from 10,000 random extractions from the posterior, shown in the left panel of Supplementary Figure 6.

$\beta(\tau)$ and the renewal equation

In an epidemic which is growing exponentially, in a deterministic manner, driven by human-to-human[1] transmission, the incidence $I(t)$ can be described by the renewal equation:

$$I(t) = \int_0^\infty I(t-\tau)\beta(\tau)d\tau, \tag{7}$$

In words, Equation 7 says that the incidence now is set by the rate at which people were infected at all previous times, weighted by how infectious those people are now. $\beta(\tau)$ is the mean rate at which an individual infects others a time τ after being infected itself. Here we take $\beta(\tau)$ to be independent of the stage of the epidemic (calendar time t): we neglect depletion of susceptible individuals through acquired immunity, changing contact patterns etc. over the timescale of the data informing our estimations of $\beta(\tau)$. After $\beta(\tau)$ has been determined, we may consider how to change it through interventions to reduce infectiousness. If one's direct and indirect infectiousness is certain to be zero after having been infected for a time T say, we only need consider the previous time window T of the epidemic – replacing the upper limit of the integral in Equation 7 by T for convenience. We take T to be infinite for generality, with $\beta(\tau)$ tending to zero at large times. Substituting into Equation 7 an exponentially growing incidence, $I(t) = I_0 e^{rt}$, gives the condition

$$1 = \int_0^\infty e^{-r\tau}\beta(\tau)d\tau, \tag{8}$$

$\beta(\tau)$ can be written as the product of two things: R_0 and the unit-normalised function $w(\tau)$

$$\beta(\tau) = R_0 w(\tau), \quad \text{with} \quad R_0 = \int_0^\infty \beta(\tau)d\tau \tag{9}$$

[1]Equation 7 describes only human-to-human transmission (though this may be indirectly via the environment, as we clarify). Excluding vector-borne diseases, zoonosis events do not scale with the number of people currently infected, and so become a negligible contribution after human-to-human transmission has begun driving exponential growth.

$w(\tau)$ is the generation time distribution – the probability density function for the time between an individual becoming infected and their subsequent onward transmission events. R_0 is the basic reproduction number. If the exponential growth rate r and the generation time distribution $w(\tau)$ have been estimated, R_0 is determined by Equation 8, i.e.

$$R_0 = 1/\int_0^\infty e^{-r\tau} w(\tau) d\tau, \tag{10}$$

We decompose $\beta(\tau)$ without any loss of generality into the following distinct contributions:

- Direct transmissions from asymptomatic individuals – those who never develop symptoms. The degree to which individuals show symptoms is of course a continuum, but a threshold can be defined for clinical purposes (i.e. sub-clinical and clinical infections) or for epidemiological purposes. We define P_a as the proportion of such individuals among all infected individuals, and $\beta_a(\tau)$ as their mean infectiousness at age-of-infection τ.

- Direct transmissions from pre-symptomatic individuals (currently without symptoms, but who will develop symptoms later). We define $\beta_p(\tau)$ as the mean infectiousness of these individuals at age-of-infection τ, conditional upon their being pre-symptomatic, which has probability $1 - s(\tau)$ where $s(\tau)$ is the cumulative distribution function of the incubation period distribution.

- Direct transmissions from symptomatic individuals (including those who have stopped showing symptoms, in general, if infectiousness may outlast symptoms), with infectiousness $\beta_s(\tau)$ conditional on having started symptoms.

- Indirect transmission via the environment. We define $\beta_e(\tau)$ as the mean rate of contaminating one's environment (with the mean being over all asymptomatic, pre-symptomatic and symptomatic individuals at age-of-infection τ). Let $E(l)$ be the rate at which contaminated environment infects new individuals a time lag l after having been contaminated. The environmentally mediated infectiousness of an individual infected a time τ ago is given by the total effect of their previous environmental contamination now: $\int_{l=0}^{\tau} \beta_e(\tau - l) E(l) d\tau$.

We therefore have, in general,

$$\beta(\tau) = P_a\beta_a(\tau) + (1 - P_a)(1 - s(\tau))\beta_p(\tau) + (1 - P_a)s(\tau)\beta_s(\tau) + \int_{l=0}^{\tau} \beta_e(\tau - l)E(l)dl \tag{11}$$

Integrating each of these terms separately gives their respective contribution to R_0:

$$R_0 = R_a + R_p + R_s + R_e \tag{12}$$

We make the following simplifying assumptions about the contributions described above, compared to the general case.

- Asymptomatic individuals are assumed to have an infectiousness proportional to that of symptomatic individuals: $\beta_a(\tau) = x_a\beta_s(\tau)$.

- Pre-symptomatic individuals are assumed to have an infectiousness equal to that of symptomatic individuals at the same age of infection: $\beta_p(\tau) = \beta_s(\tau)$

- The rate at which individuals contaminate their environment, $\beta_e(\tau)$ is assumed to be proportional to the direct infectiousness of symptomatic individuals $\beta_s(\tau)$. The proportionality constant can be absorbed into the function $E(l)$ which multiplies $\beta_e(\tau)$, so that we have $\beta_e(\tau) = \beta_s(\tau)$.

The $\beta(\tau)$ we consider is therefore

$$\beta(\tau) = \underbrace{P_a x_a \beta_s(\tau)}_{\text{asymptomatic}} + \underbrace{(1-P_a)(1-s(\tau))\beta_s(\tau)}_{\text{pre-symptomatic}} + \underbrace{(1-P_a)s(\tau)\beta_s(\tau)}_{\text{symptomatic}} + \underbrace{\int_{l=0}^{\tau} \beta_s(\tau-l)E(l)dl}_{\text{environmental}} \quad (13)$$

We solved for the form of $\beta(\tau)$ above first by fitting the shape of the pre-symptomatic plus symptomatic contributions to our inferred generation time interval: these functions are proportional to each other when the transmission pairs analysed for the generation time distribution represent pre-symptomatic and symptomatic exposure in the proportion representative of overall epidemic spread. We make that assumption here. This assumption would be violated by biased selection of transmission pairs for sampling. For example if the infector being in a later, symptomatic stage of infection makes identification of the pair more likely, then a data set of identified pairs will be undersampled for pre-symptomatic exposure and will overestimate the typical generation time. The next step in solving the model was calculating the relative scaling constant of the environmental contribution to $\beta(\tau)$ to give the required R_E/R_0, and finally the overall scaling constant of $\beta(\tau)$ is determined to reproduce the observed exponential growth rate.

To determine uncertainty, we drew 10,000 input parameter sets from the uncertainties shown in main text Table 2. For the data-driven parameters, these uncertainties are likelihoods, which can be interpreted as posteriors if one's prior is an improper uniform distribution; we fit lognormal distributions to the 95% CIs and central estimates in order to obtain the full distribution. For the other parameters, the uncertainty distributions are pure priors.

Derivation of the impact of interventions

Impact of interventions. To calculate the impact of contact tracing and isolation, we followed the mathematical treatment of (9), explained in detail in the Supplementary Information. Specifically, we calculated the epidemic dynamics of the quantity Y(t, , '): the number of individuals at time t who were infected at a time t - by individuals who were in turn infected at time t - ', subject to case isolation and contact tracing interventions. Both interventions are assumed to be immediate upon individuals showing symptoms, but both have efficacies that can vary continuously between 0 and 1.

One difference in notation is that we use $s(\tau)$ here to denote the probability of having started showing symptoms, corresponding to their $(1 - S(\tau))$.

Prelude: describing epidemic growth with no intervention

Let $Y(t, \tau, \tau')$ be the number[2] of individuals at time t who were infected at a time $t - \tau$ by individuals who were in turn infected at a time $t - \tau'$. Y satisfies $Y(t + dt, \tau + dt, \tau' + dt) = Y(t, \tau, \tau')$, 'translational invariance', because incrementing all three arguments by exactly the same value means following exactly the same cohort of individuals through to a different moment in time. Equivalently,

$$\frac{\partial Y(t,\tau,\tau')}{\partial t} + \frac{\partial Y(t,\tau,\tau')}{\partial \tau} + \frac{\partial Y(t,\tau,\tau')}{\partial \tau'} = 0 \quad (14)$$

In the absence of any intervention, $Y(t, \tau, \tau')$ satisfies the generalised Kermack-McKendrick equations (also referred to as the Von Foerster equations):

[2]Technically Y is a double-density: one must integrate it over a range of τ values and a range of τ' values to get an actual number of people at time t.

$$Y(t,0,\tau) = \beta(\tau) \int_{\tau'=\tau}^{\infty} Y(t,\tau,\tau')d\tau' \tag{15}$$

In words, equation 15 says that the incidence of newly infected individuals at time t due to individuals who were themselves infected a time τ ago is given by the current infectiousness of those individuals infected a time τ ago, multiplied by how many of those individuals there are: integrating their number over all possible times that their infector was infected. Since the number of infected individuals is self-renewing, we anticipate exponential dependence on t. Together with translational invariance, this implies the general form

$$Y(t,\tau,\tau') = y(\tau'-\tau)e^{r(t-\tau)} \tag{16}$$

Substituting equation 16 into equation 15 gives

$$y(\tau)e^{rt} = \beta(\tau) \int_{\tau'=\tau}^{\infty} y(\tau'-\tau)e^{r(t-\tau)}d\tau' \tag{17}$$

$$y(\tau) = e^{-r\tau}\beta(\tau) \int_{\tau'=\tau}^{\infty} y(\tau'-\tau)d\tau' \tag{18}$$

$$= e^{-r\tau}\beta(\tau) \int_{\rho=0}^{\infty} y(\rho)d\rho \tag{19}$$

The solution to equation 19 is, for any value of the constant y_0,

$$y(\tau) = y_0 e^{-r\tau}\beta(\tau), \quad \text{with the constraint } \int_{\rho=0}^{\infty} e^{-r\rho}\beta(\rho)d\rho = 1 \tag{20}$$

Decomposing $\beta(\tau)$ into two factors – its integral R_0, and the unit-normalised generation time interval $w(\tau)$ – the constraint above gives the relationship between r, R_0 and $w(\tau)$. Substituting this solution for $y(\tau)$ back into equation 16 gives

$$Y(t,\tau,\tau') = y_0 e^{-r(\tau'-\tau)}\beta(\tau'-\tau)e^{r(t-\tau)} \tag{21}$$

$$= y_0 e^{r(t-\tau')}\beta(\tau'-\tau) \tag{22}$$

The impact of case isolation, contact tracing and quarantine

In the presence of case isolation of efficacy ϵ_I and contact tracing plus quarantine of efficacy ϵ_T, equation 15 is modified (see Fraser, Riley et al. 2004) to

$$Y(t,0,\tau) = \beta(\tau)\left(1-\epsilon_I s(\tau)\right) \int_{\tau}^{\infty} \left(1-\epsilon_T+\epsilon_T\frac{1-s(\tau')}{1-s(\tau'-\tau)}\right) Y(t,\tau,\tau')d\tau' \tag{23}$$

Note that we take the upper limit of the integral here to be ∞ rather than t (so that the epidemic is solved exactly by exponential growth, instead of beginning according to specific boundary conditions and then tending toward exponential growth as the boundary conditions are forgotten). Translational invariance together with exponential growth with t imply the same general form as previously – equation 16 – but with a different functional form for $y(\tau)$. Substituting equation 16 into equation 23 gives the 'next-generation' equation for y:

$$y(\tau) = e^{-r\tau}\beta(\tau)\left(1-\epsilon_I s(\tau)\right) \int_{\tau'=\tau}^{\infty} \left(1-\epsilon_T+\epsilon_T\frac{1-s(\tau')}{1-s(\tau'-\tau)}\right) y(\tau'-\tau)d\tau' \tag{24}$$

$$= e^{-r\tau}\beta(\tau)\left(1-\epsilon_I s(\tau)\right) \int_{\rho=0}^{\infty} \left(1-\epsilon_T\frac{s(\rho+\tau)-s(\rho)}{1-s(\rho)}\right) y(\rho)d\rho \tag{25}$$

redefining the integration variable to be $\rho = \tau' - \tau$ for convenience. Hence, the growth rate after the interventions corresponds to the value of r for which the functional linear 'next-generation' operator

$$\mathcal{N}_r y = e^{-r\tau}\beta(\tau)\left(1-\epsilon_I s(\tau)\right)\int_0^\infty \left(1-\epsilon_T \frac{s(\rho+\tau)-s(\rho)}{1-s(\rho)}\right) y(\rho)d\rho \tag{26}$$

has the largest eigenvalue equal to 1. In other words, given the eigenvalue equation $\mathcal{N}_r y = \nu_r^{max} y$, r is determined by $\nu_r^{max} = 1$, in a functional generalization of the Euler-Lotka equation.

If R_0 is unknown, but the generation time distribution $\omega(\tau) = \beta(\tau)/R_0$ is known, then we consider the operator

$$\mathcal{O} = \mathcal{N}_0/R_0 = \omega(\tau)\left(1-\epsilon_I s(\tau)\right)\int_0^\infty \left(1-\epsilon_T \frac{s(\rho+\tau)-s(\rho)}{1-s(\rho)}\right)(\cdot)d\rho \tag{27}$$

By construction, this operator has maximal eigenvalue $1/R_0$ if and only if the value of R_0 corresponds to a growth rate $r = 0$ after interventions. Then, the inverse of the largest eigenvalue of $\mathcal{O}$ is precisely the maximum value of R_0 for which $R \leq 1$ in the presence of interventions.

This approach includes potential transmission from a fraction P_a of completely asymptomatic individuals. Denote by x_a the relative infectiousness of those individuals (which could be different from 1). For the purpose of this model, this is effectively equivalent to having $P_a x_a$ infectious individuals with relative infectiousness 1. Hence, the effective fraction of asymptomatic individuals is $P_a x_a/(1 - P_a + P_a x_a)$. Then, the fraction of infected individuals that have already shown symptoms a time τ after infection corresponds to

$$s(\tau) = \frac{1-P_a}{1-P_a+P_a x_a}\int_0^\tau i(\tau')d\tau' \tag{28}$$

where $i(\tau)$ is the incubation time distribution for individuals that will eventually become symptomatic.

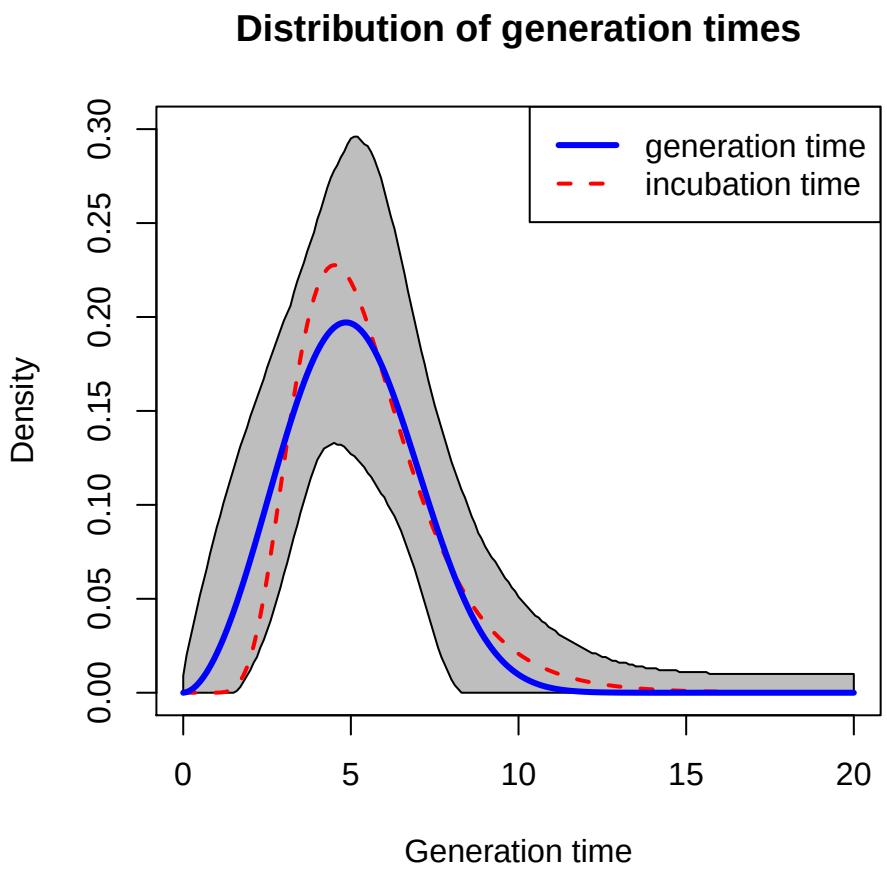
This approach can be modified to include environmental transmission as well. The timing of environmentally mediated transmission is shifted to slightly later in the infection (larger τ) due to the waiting time of the virus in the environment in between hosts, i.e. due to the function $E(l)$ discussed earlier not being a delta function concentrated at zero. However as an approximation for analytical tractability, to derive the effect of the intervention we neglect this shift in timing i.e. we neglect the viral waiting time in the environment compared to the range of τ values over which $\beta(\tau)$ is spread. With this approximation, environmental transmission events can be treated as cases of untraceable transmission. Isolation and quarantine are effective in preventing environmental spread, since they stop the infected individual from releasing the virus in the environment. (The only relevant exception is environment transmission within isolation facilities, which is not prevented by isolation - quite the opposite.) We assume here that all environmental transmission events correspond simply to untraceable transmissions, and that the rate of viral shedding in the environment is proportional to the person-to-person infectiousness. Then, the theory developed above can be applied by replacing the efficacy of contact tracing $\epsilon_T \to \epsilon_T(1 - R_e/R_0)$, where the factor $1 - R_e/R_0$ corresponds to the fraction of traceable transmissions and ϵ_T is the efficacy per traceable transmission. Hence, the results above can be generalised to include environmental transmission by considering the next-generation operator

$$\mathcal{N}_r = e^{-r\tau}\beta(\tau)\left(1-\epsilon_I s(\tau)\right)\int_0^\infty \left(1-\epsilon_T\left(1-\frac{R_e}{R_0}\right)\frac{s(\rho+\tau)-s(\rho)}{1-s(\rho)}\right)(\cdot)d\rho \tag{29}$$

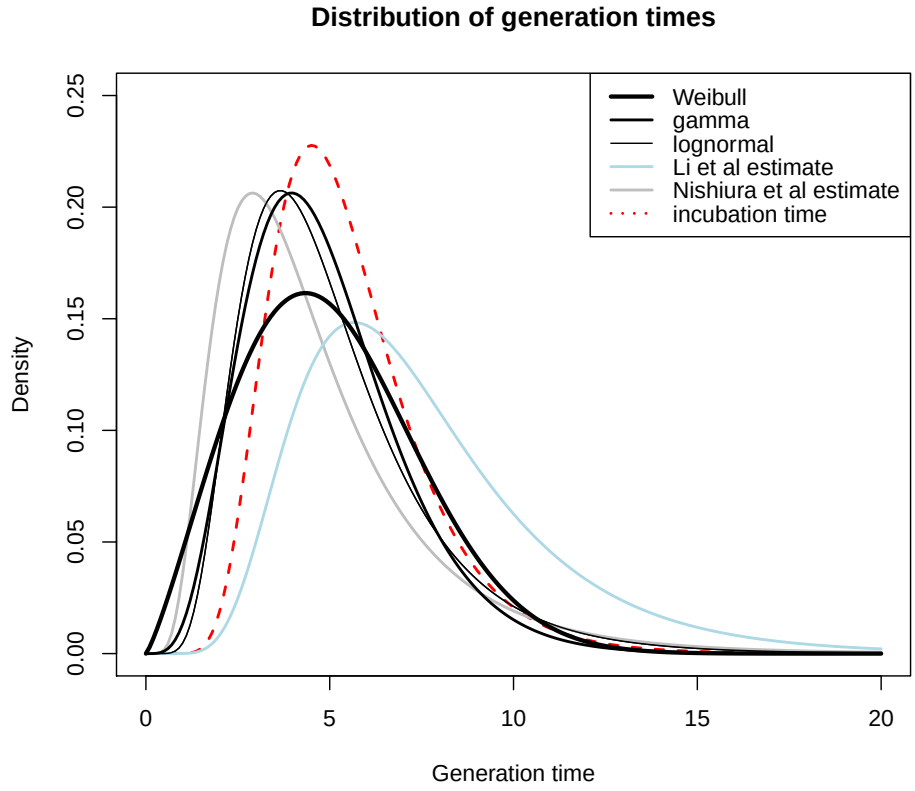
Finally, a delay Δ in both isolation of the index case and contact tracing/quarantine of contacts can be modelled simply by replacing all occurrences of $s(\tau)$ with $s(\tau - \Delta)$.

Numerically, the eigenfunction corresponding to the largest eigenvalue can be found by iterating the operator $\mathcal{N}_r$. Up to a constant factor, the exact solution is given by $\lim_{k\to\infty} \mathcal{N}_r^k y_0$ where y_0 is an arbitrary initial condition; in practice, $\mathcal{N}_r^k y_0$ for large enough k provides a good approximation to the solution, with exponentially fast convergence (at a rate given by the ratio of second- and first-most dominant eigenvalues per application of the operator). We found that fewer than 10 iterations provided an excellent approximation in the parameter space and for the distributions and values of R_0 considered in this paper.

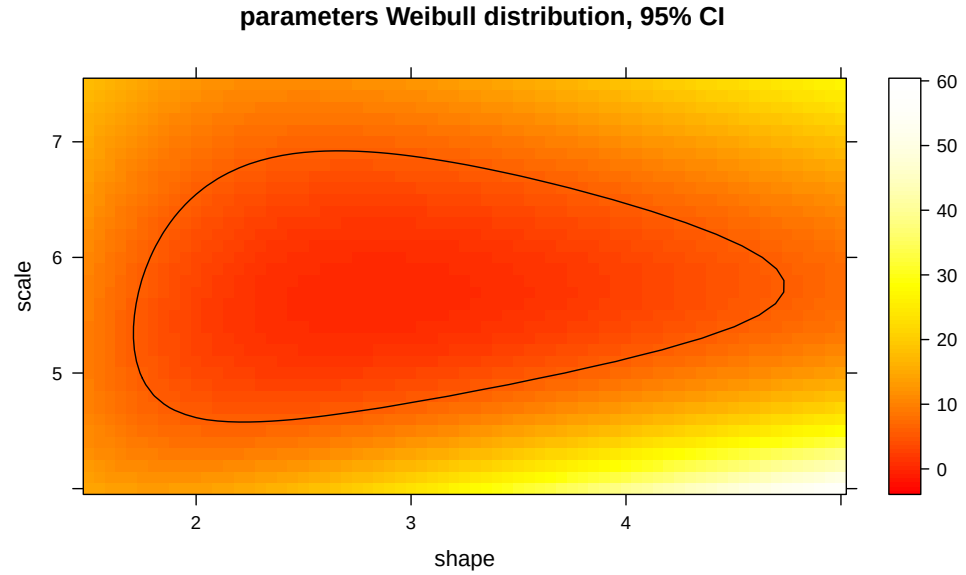
Supplementary Figures



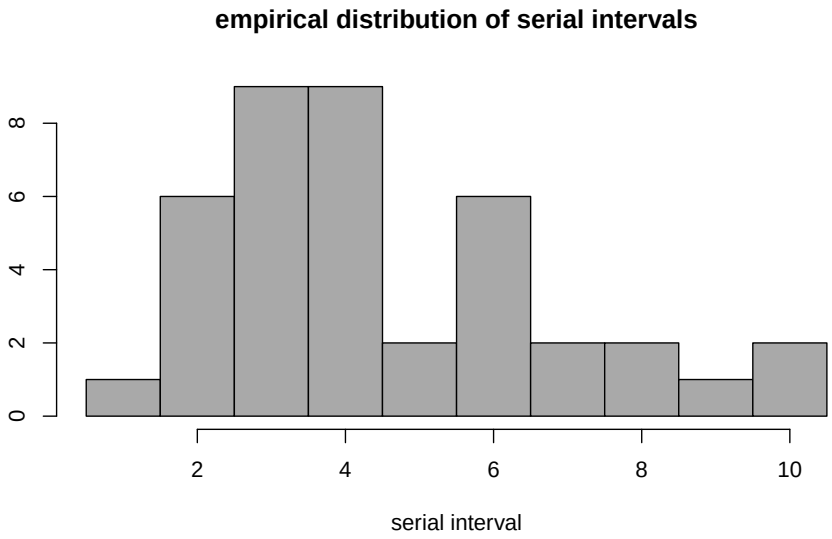
Supplementary Figure 1: Same as Figure 1 left panel in Main Text, showing only the best-fit Weibull distribution for the generation time and the 95% CI of the point estimate.



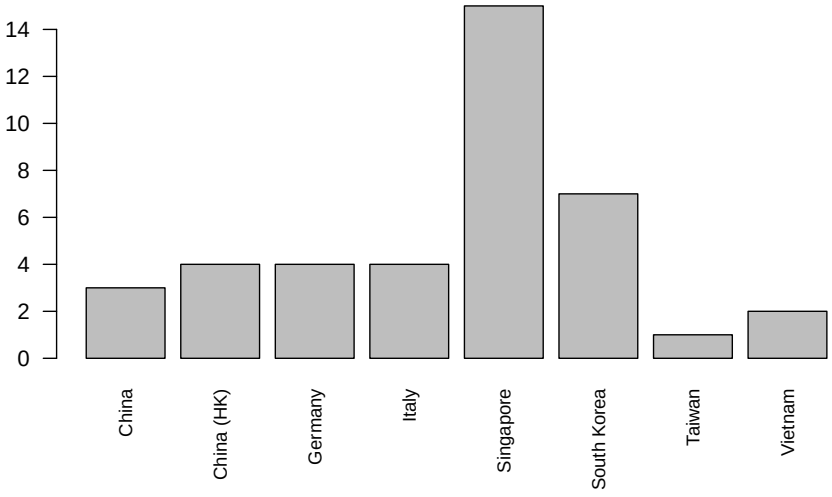
Supplementary Figure 2: Same as Figure 1 left panel in Main Text, but the generation time distribution is inferred from the likelihood function conditional on the occurrence of the transmission events.



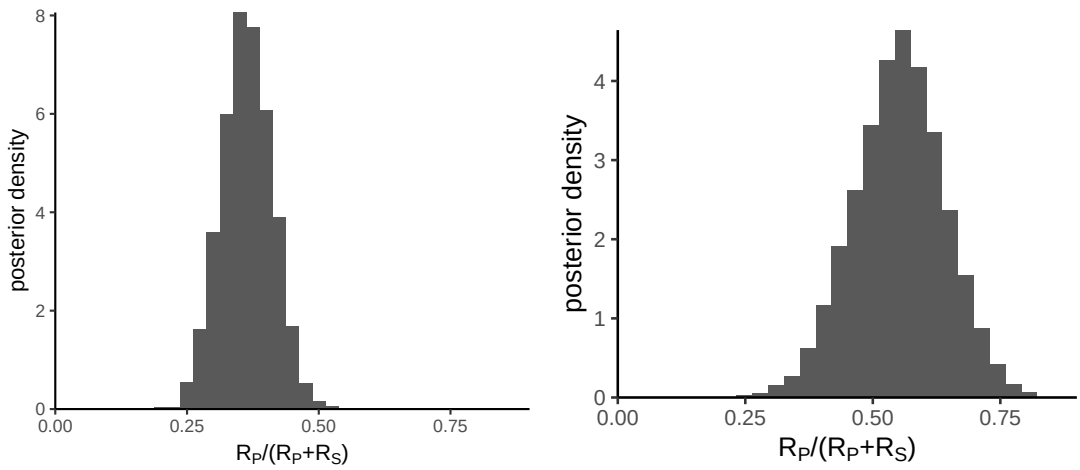
Supplementary Figure 3: Contour of the 95% confidence interval for the shape and scale parameters of the Weibull distribution of generation times. The colours represent the $-2\Delta \log \mathcal{L}$ test statistic (lower values correspond to parameters for which the observed data is more likely).



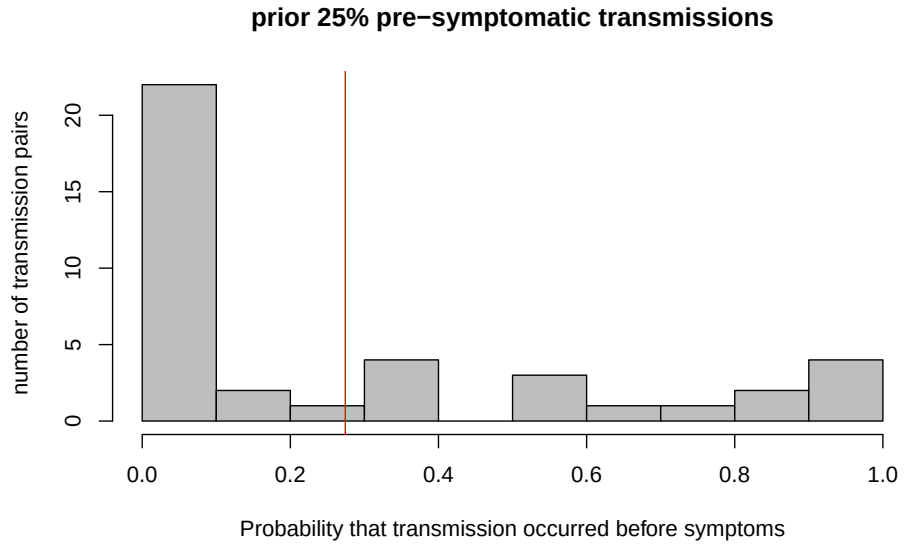
Supplementary Figure 4: Distribution of serial intervals (i.e. time from onset of symptoms in source to onset of symptoms in recipient) from the 40 transmission pairs analysed in this study.



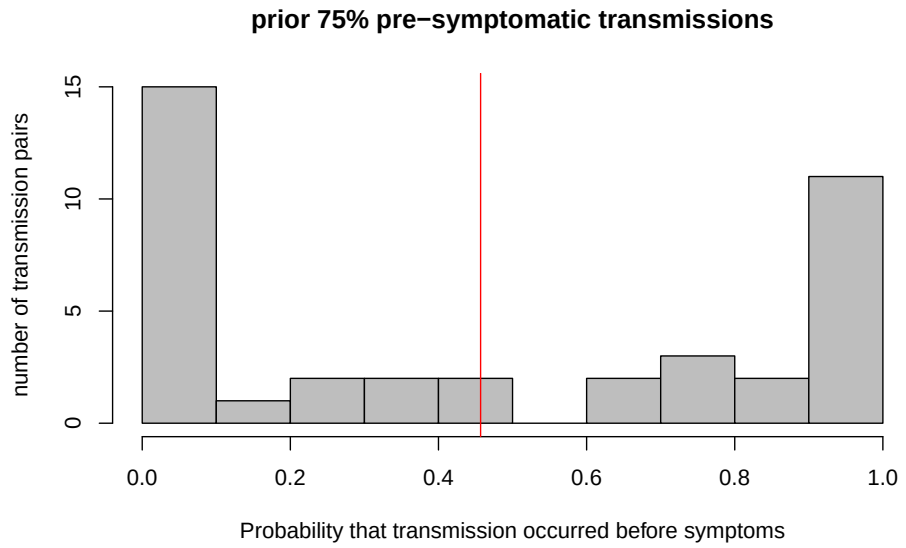
Supplementary Figure 5: Geographical distribution of transmission occurrence among the 40 source-recipient pairs analysed in this study.



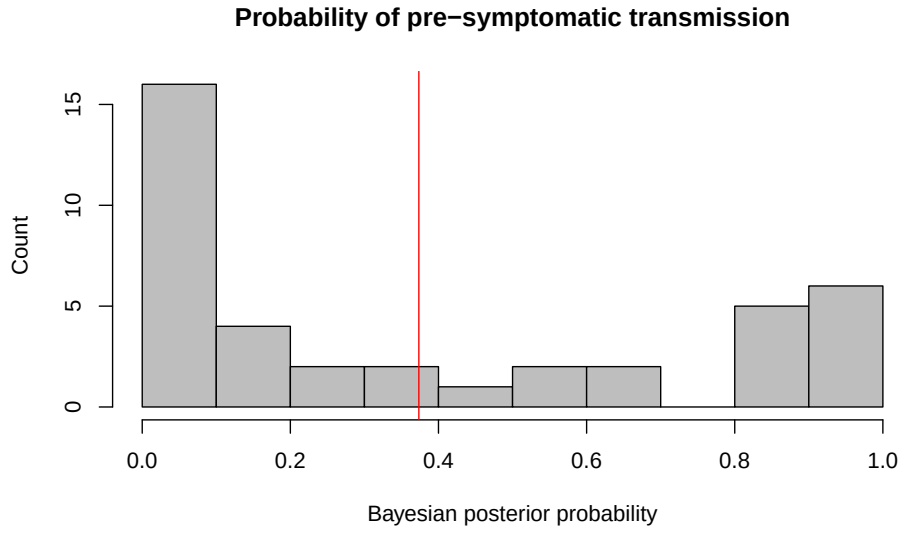
Supplementary Figure 6: Uncertainty on the contribution of pre-symptomatic transmissions to all transmissions from infected individuals that will eventually show symptoms. Left: posterior distribution from the 40 transmission pairs analysed in this study. Right: posterior distribution from the infectiousness model approach.



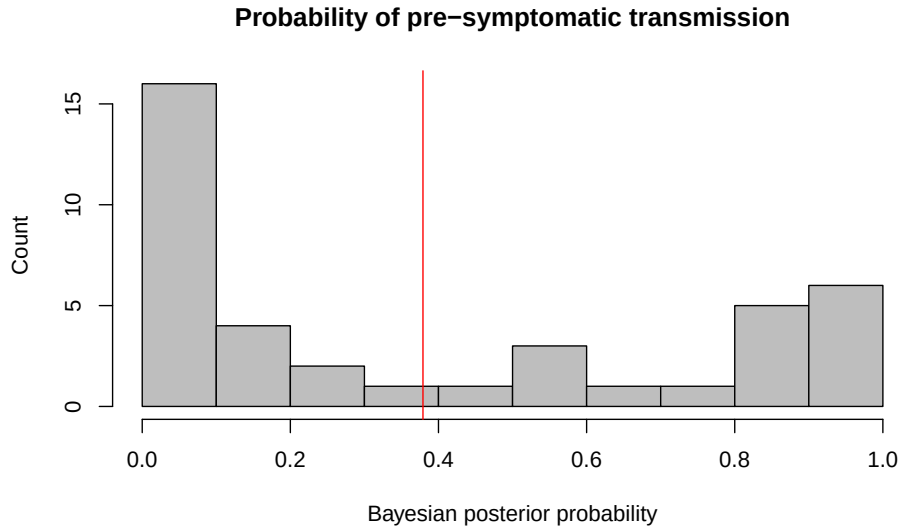
Supplementary Figure 7: Same as Figure 1 right panel in Main Text, but assuming a 25% prior for pre-symptomatic transmission.



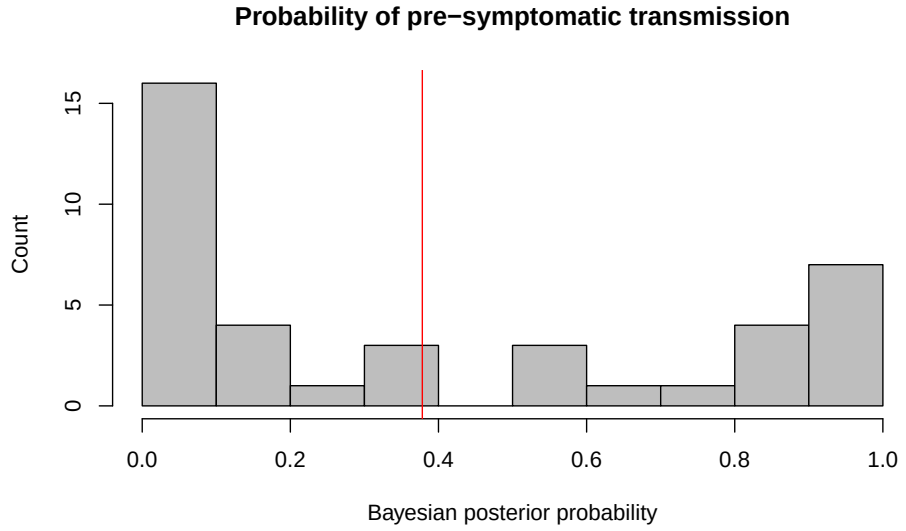
Supplementary Figure 8: Same as Figure 1 right panel in Main Text, but assuming a 75% prior for pre-symptomatic transmission.



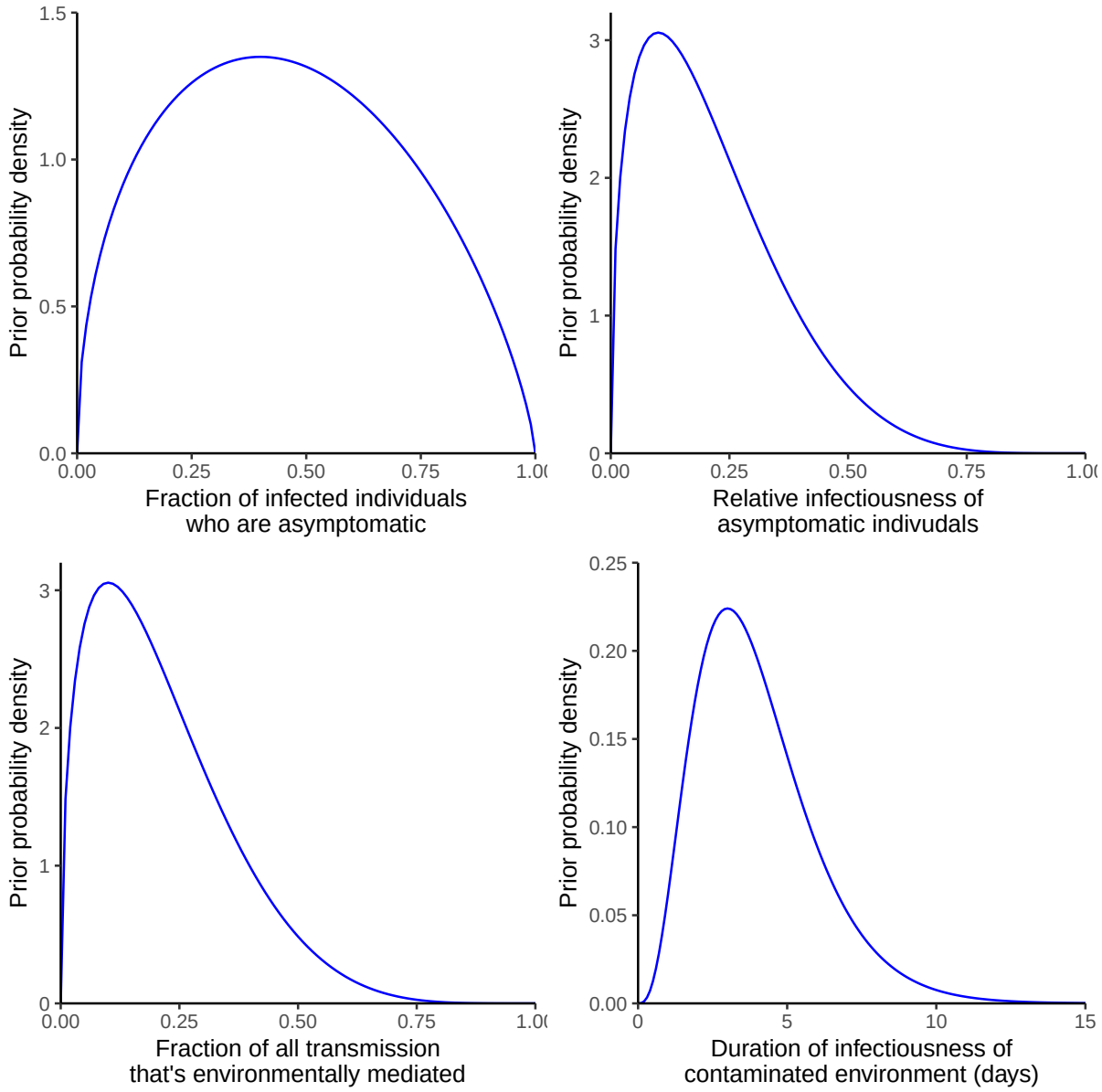
Supplementary Figure 9: Same as Figure 1 right panel in Main Text, but assuming a gamma distribution of generation times.



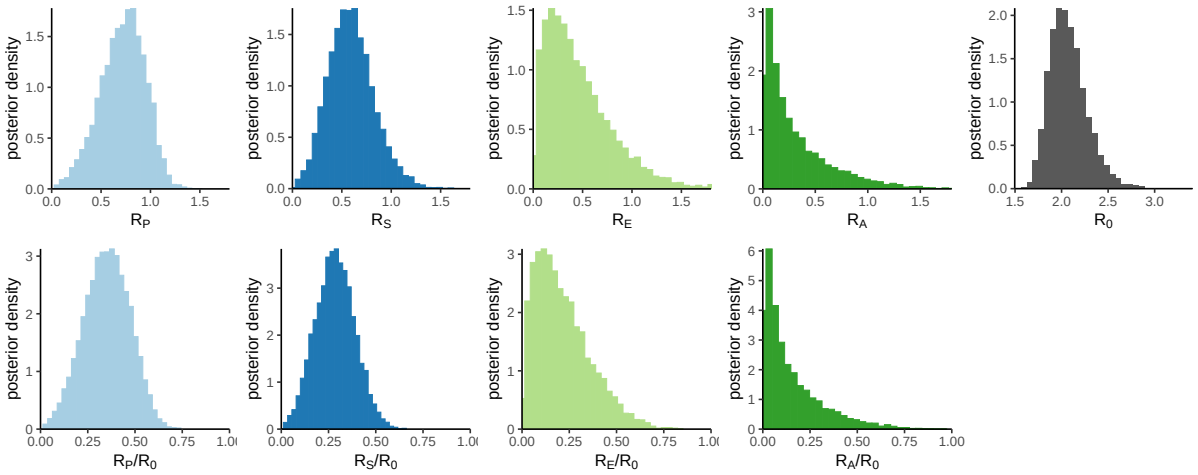
Supplementary Figure 10: Same as Figure 1 right panel in Main Text, but assuming a lognormal distribution of generation times.



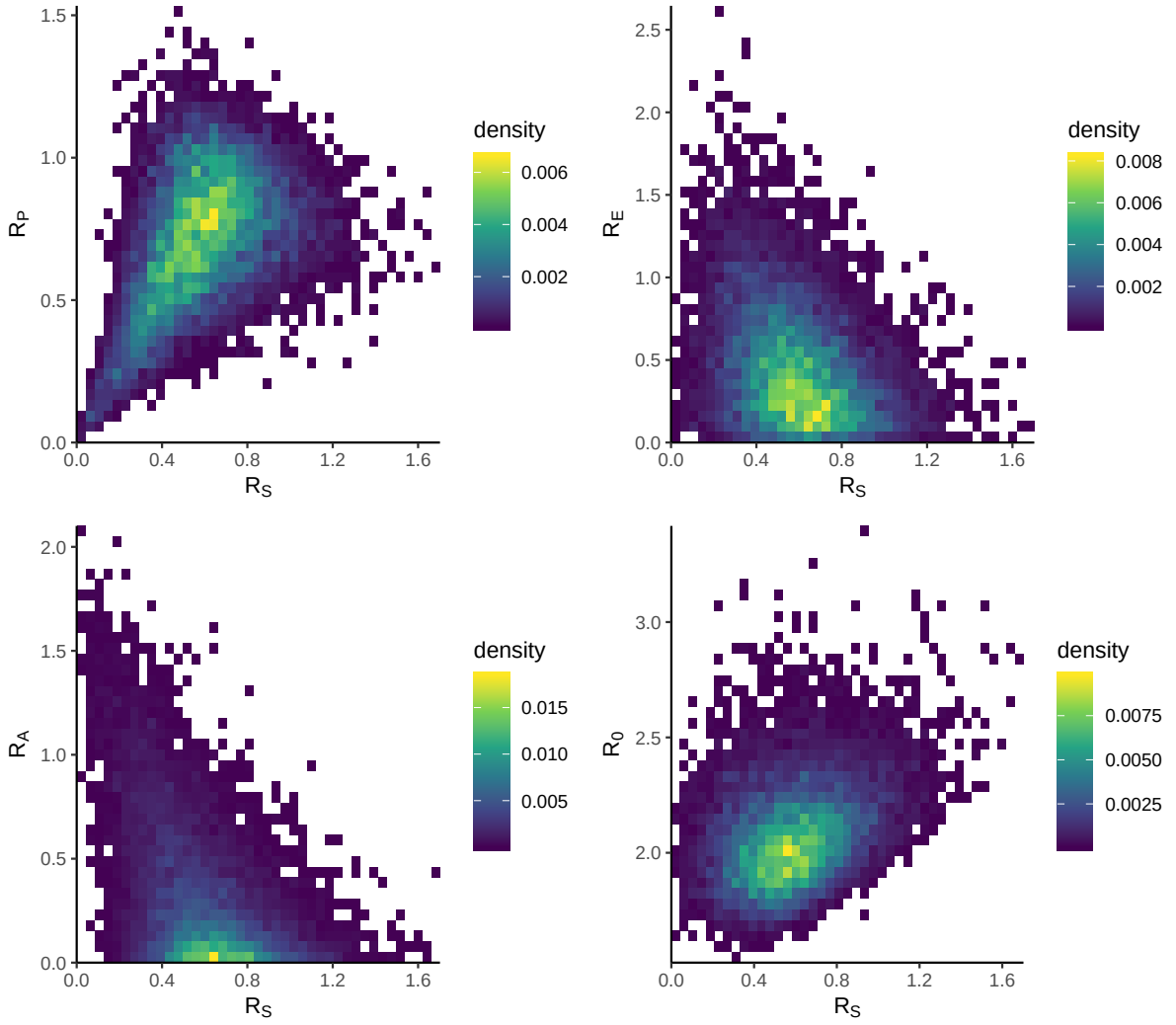
Supplementary Figure 11: Same as Figure 1 right panel in Main Text, but the posterior probabilities of pre-symptomatic transmission are inferred from the likelihood function conditional on the occurrence of the transmission events.



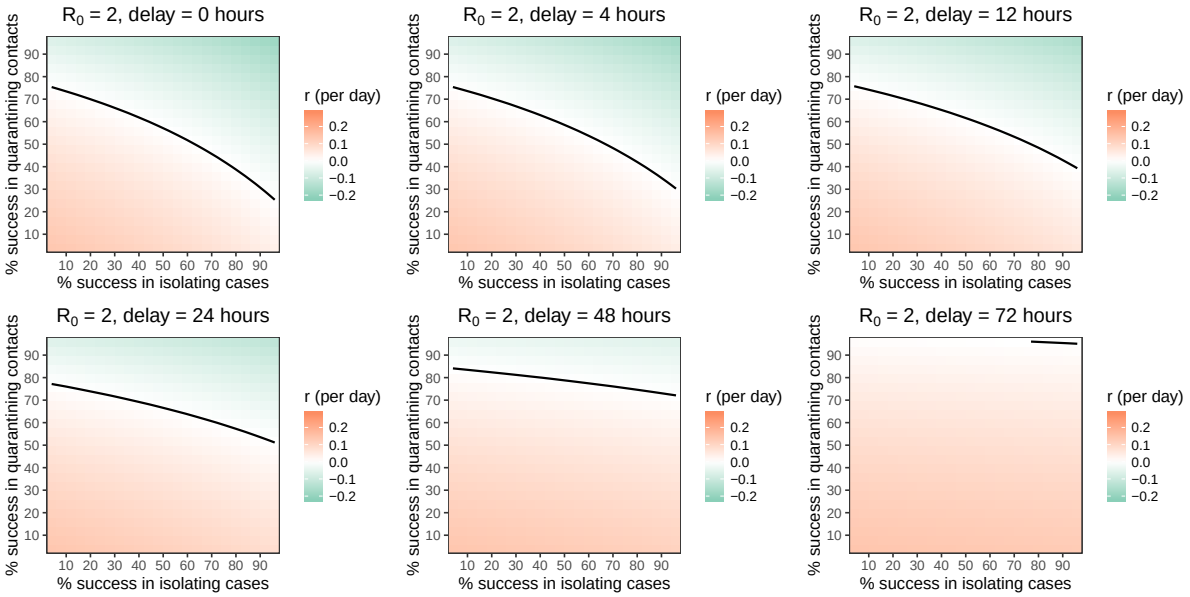
Supplementary Figure 12: The prior probability density functions we chose for the parameters in Table 1 of the main text.



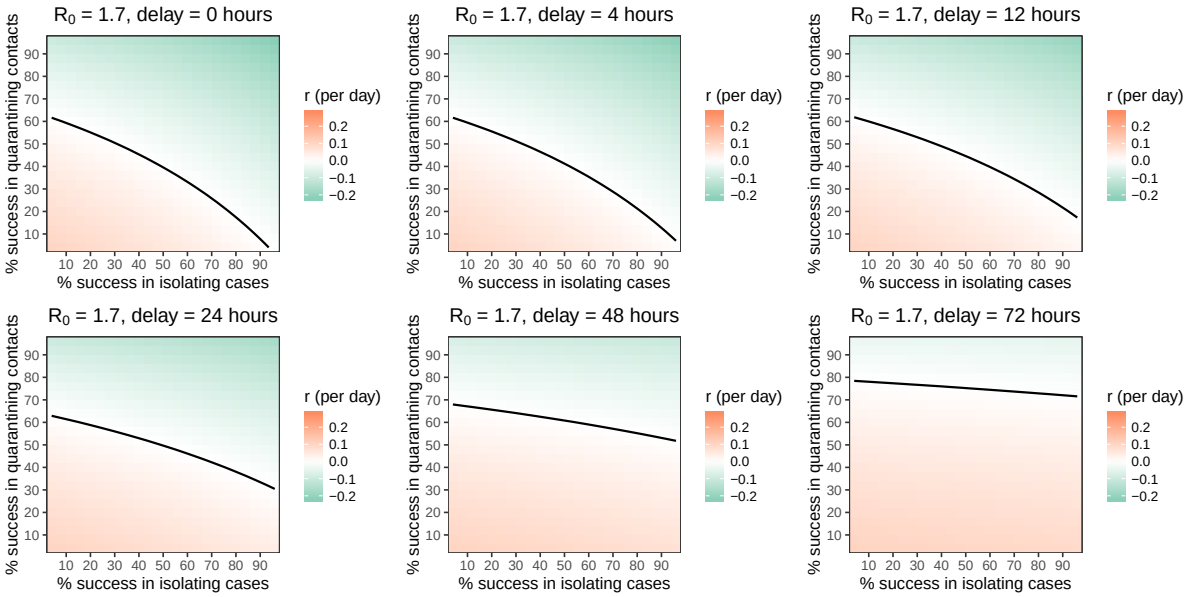
Supplementary Figure 13: The probability density for output parameters of the infectiousness model, from sampling the uncertainty distributions of the input parameters. The top panels show absolute reproduction numbers: the four contributors to R_0 (from pre-symptomatic, symptomatic, environmental and asymptomatic transmission respectively, with the bin at 1.8 containing overflow) and R_0 itself. The bottom panels show the fractions of R_0 that each contribution represents. The colours match those of main text Figure 2.



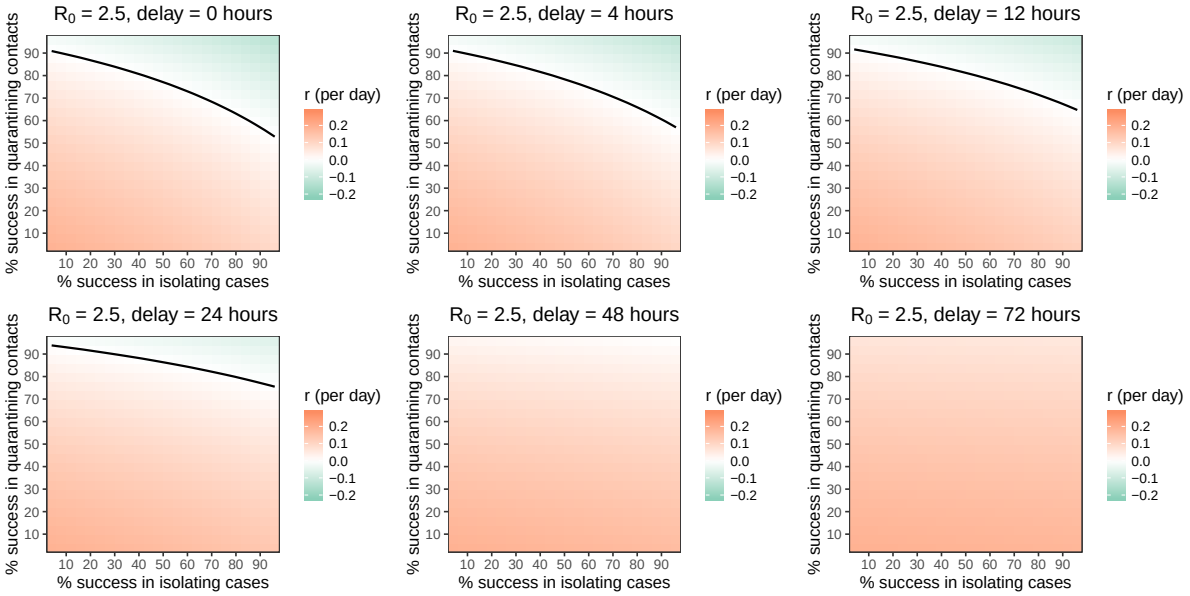
Supplementary Figure 14: Showing the correlations in the uncertainties between the four different contributions to R_0, and R_0 itself (i.e. the correlations between the results of Figure 3 from the main text). R_S is correlated with R_P because their ratio is relatively well constrained by the measured $w(\tau)$ and incubation period distributions. Both of these are anti-correlated with the R_A and R_S at the tails of the distributions because the sum of all four contributions is constrained (to R_0) by the observed r and $w(\tau)$. There is little correlation between R_S and the overall R_0.



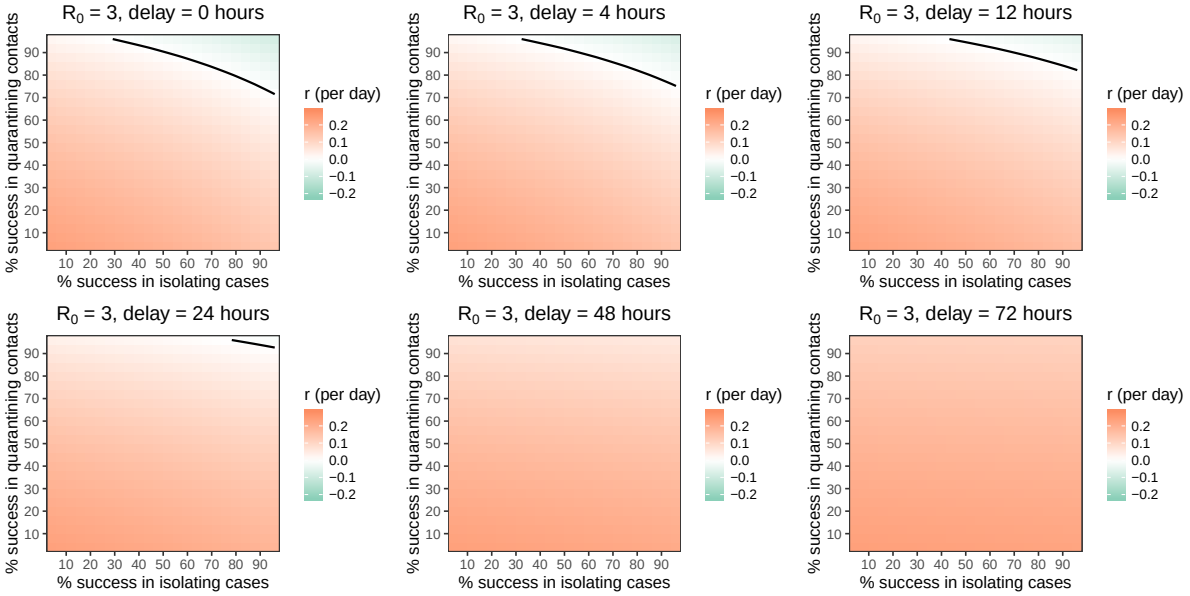
Supplementary Figure 15: As Figure 4 in the main text, but instead of showing uncertainty in the position of $r = 0$ due to uncertainty in R_0, as in main text Figure 4, we show results for each different value of R_0 separately in different plots: see Supplementary Figures 16 – 19. This plot shows the central estimate of R_0, namely 2, and the position of the solid black line (indicating the threshold for epidemic control, at $r = 0$) is in the same location as in main text Figure 4. The fraction of transmissions that are environmentally mediated is fixed at 10%, and the fraction from asymptomatic individuals at 6.25% (corresponding to our central estimates of $P_a = 40\%$, $x_a = 0.1$), as in the main text.



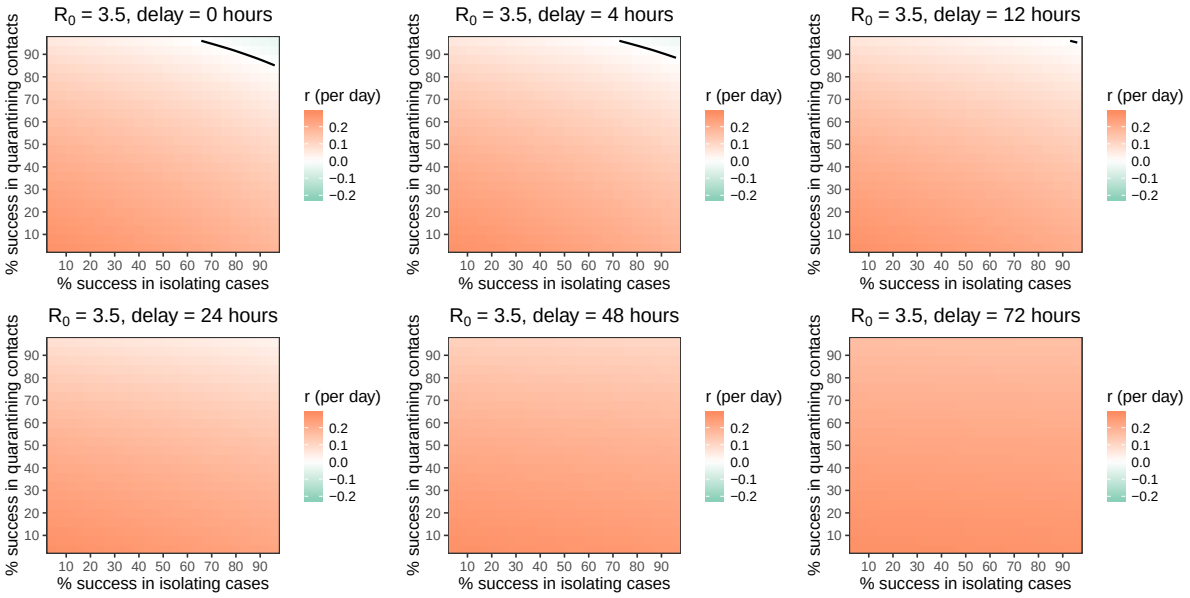
Supplementary Figure 16: As Supplementary Figure 15, but using the lower bound R_0 value of 1.7 instead of the central estimate of 2. The position of the solid black line here corresponds to the position of the innermost dashed black line in main text Figure 4, i.e. it indicates the uncertainty in where the $r = 0$ line might be in the intervention parameter space (x and y axes) due to uncertainty in the value of R_0.



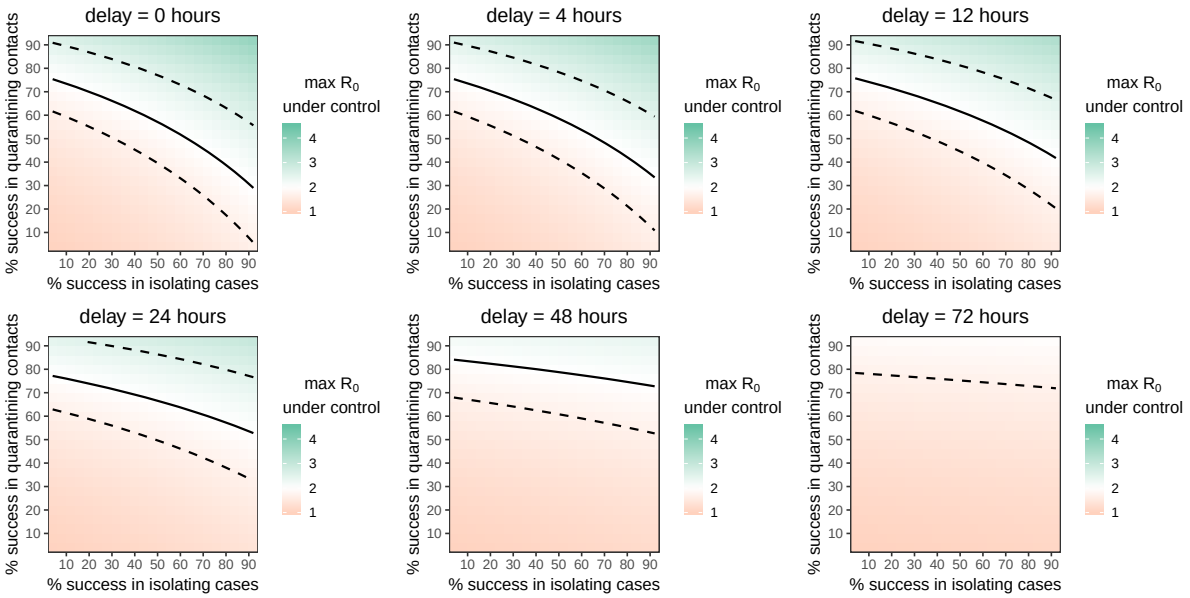
Supplementary Figure 17: As Supplementary Figure 15, but using the upper bound R_0 value of 2.5 instead of the central estimate of 2. The position of the solid black line here corresponds to the position of the outermost dashed black line in main text Figure 4, i.e. it indicates the uncertainty in where the $r = 0$ line might be in the intervention parameter space (x and y axes) due to uncertainty in the value of R_0.



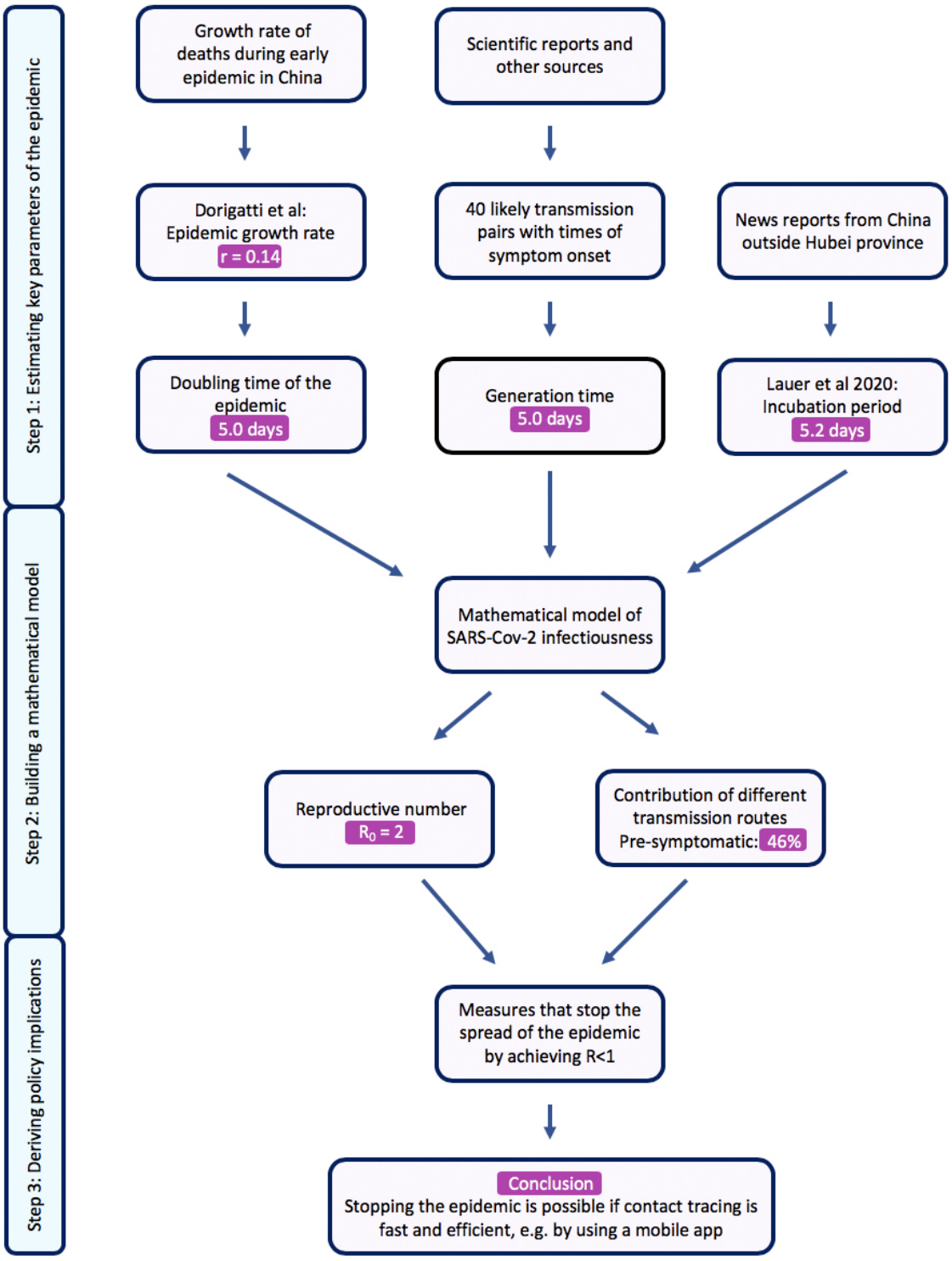
Supplementary Figure 18: As Supplementary Figure 15, but using the upper bound R_0 value of 3 instead of the central estimate of 2. Such value could be more appropriate to describe the epidemiological dynamics in Europe. This plot shows how high success rates, more sophisticated tracking algorithms or some degree of physical distancing could be required in European countries.



Supplementary Figure 19: As Supplementary Figure 15, but using the upper bound R_0 value of 3.5 instead of the central estimate of 2. Such value could be more appropriate to describe the epidemiological dynamics in Europe. This plot shows how high success rates, more sophisticated tracking algorithms or some degree of physical distancing could be required in European countries.



Supplementary Figure 20: Heat map plot showing the maximum possible R_0 that could be contained (by reduction to less than 1) for a pathogen with the same generation time distribution that we have inferred for SARS-CoV-2, as a function of the success rate of isolation of symptomatic cases (x axis) and the success rate of contact tracing (y axis). Contours of constant R_0 show the different combinations of the two success rates that would be able to control the epidemic for an R_0 of that value. The solid black line is such a contour for our central estimate of R_0 for SARS-CoV-2, namely $R_0 = 2$; this line therefore shows the threshold for epidemic control. The dashed lines show the contours of the upper and lower bounds of the CI for R_0 (1.7 - 2.5), i.e. these indicate one source of uncertainty in the epidemic control threshold. Each of the six panels shows results for a different delay from initiation of symptoms to isolation of the case and quarantine of their contacts (the same delay for both). The fraction of transmissions that are environmentally mediated is fixed at 10%, and the fraction from asymptomatic individuals at 6.25% (corresponding to our central estimates of $P_a = 40\%$, $x_a = 0.1$), as in the main text.



Supplementary Figure 21: A graphical summary of our analysis.

References

1. World Health Organization, *Coronavirus Disease 2019 (COVID-19): Situation Report – 36* (25 February 2020); www.who.int/docs/default-source/coronaviruse/situation-reports/20200225-sitrep-36-covid-19.pdf.

2. W.-J. Guan, Z.-Y. Ni, Y. Hu, W.-H. Liang, C.-Q. Ou, J.-X. He, L. Liu, H. Shan, C.-L. Lei, D. S. C. Hui, B. Du, L.-J. Li, G. Zeng, K.-Y. Yuen, R.-C. Chen, C.-L. Tang, T. Wang, P.-Y. Chen, J. Xiang, S.-Y. Li, J.-L. Wang, Z.-J. Liang, Y.-X. Peng, L. Wei, Y. Liu, Y.-H. Hu, P. Peng, J.-M. Wang, J.-Y. Liu, Z. Chen, G. Li, Z.-J. Zheng, S.-Q. Qiu, J. Luo, C.-J. Ye, S.-Y. Zhu, N.-S. Zhong, China Medical Treatment Expert Group for Covid-19, Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* 10.1056/NEJMoa2002032 (2020). doi:10.1056/NEJMoa2002032 Medline

3. H. Li, S. Wang, F. Zhong, W. Bao, Y. Li, L. Liu, H. Wang, Y. He, Age-dependent risks of Incidence and Mortality of COVID-19 in Hubei Province and Other Parts of China. medRxiv 2020.02.25.20027672 [preprint]. 6 March 2020.

4. N. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, C. Walters, P. Winskill, C. Whittaker, C. Donnelly, S. Riley, A. Ghani, *Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand* (16 March 2020); www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf.

5. I. T. S. Yu, J. J. Y. Sung, The epidemiology of the outbreak of severe acute respiratory syndrome (SARS) in Hong Kong—What we do know and what we don't. *Epidemiol. Infect.* **132**, 781–786 (2004). doi:10.1017/S0950268804002614 Medline

6. Y. Li, X. Huang, I. T. S. Yu, T. W. Wong, H. Qian, Role of air distribution in SARS transmission during the largest nosocomial outbreak in Hong Kong. *Indoor Air* **15**, 83–95 (2005). doi:10.1111/j.1600-0668.2004.00317.x Medline

7. J. A. Otter, C. Donskey, S. Yezli, S. Douthwaite, S. D. Goldenberg, D. J. Weber, Transmission of SARS and MERS coronaviruses and influenza virus in healthcare settings: The possible role of dry surface contamination. *J. Hosp. Infect.* **92**, 235–250 (2016). doi:10.1016/j.jhin.2015.08.027 Medline

8. S. W. X. Ong, Y. K. Tan, P. Y. Chia, T. H. Lee, O. T. Ng, M. S. Y. Wong, K. Marimuthu, Air, Surface Environmental, and Personal Protective Equipment Contamination by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) From a Symptomatic Patient. *JAMA* 10.1001/jama.2020.3227 (2020). doi:10.1001/jama.2020.3227 Medline

9. C. Fraser, S. Riley, R. M. Anderson, N. M. Ferguson, Factors that make an infectious disease outbreak controllable. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6146–6151 (2004). doi:10.1073/pnas.0307506101 Medline

10. C. M. Peak, L. M. Childs, Y. H. Grad, C. O. Buckee, Comparing nonpharmaceutical interventions for containing emerging epidemics. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4023–4028 (2017). doi:10.1073/pnas.1616438114 Medline

11. J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, S. Funk, R. M. Eggo, F. Sun, S. Flasche, B. J. Quilty, N. Davies, Y. Liu, S. Clifford, P. Klepac, M. Jit, C. Diamond, H. Gibbs, K. van Zandvoort, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob. Health* **8**, e488–e496 (2020). doi:10.1016/S2214-109X(20)30074-7 Medline

12. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. K. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020). doi:10.1056/NEJMoa2001316 Medline

13. Z.-D. Tong, A. Tang, K.-F. Li, P. Li, H.-L. Wang, J.-P. Yi, Y.-L. Zhang, J.-B. Yan, Potential Presymptomatic Transmission of SARS-CoV-2, Zhejiang Province, China, 2020. *Emerg. Infect. Dis.* **26**, (2020). doi:10.3201/eid2605.200198 Medline

14. Y. Bai, L. Yao, T. Wei, F. Tian, D.-Y. Jin, L. Chen, M. Wang, Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA* 10.1001/jama.2020.2565 (2020). doi:10.1001/jama.2020.2565 Medline

15. Channel News Asia; https://infographics.channelnewsasia.com/covid-19/coronavirus-singapore-clusters.html?cid=FBcna.

16. Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S. A. Truelove, T. Zhang, W. Gao, C. Cheng, X. Tang, X. Wu, Y. Wu, B. Sun, S. Huang, Y. Sun, J. Zhang, T. Ma, J. Lessler, T. Feng, Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. medRxiv 2020.03.03.20028423 [preprint]. 27 March 2020.

17. K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Euro Surveill.* **25**, (2020). doi:10.2807/1560-7917.ES.2020.25.10.2000180 Medline

18. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. Walker, H. Fu, A. Dighe, J. Griffin, A. Cori, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. M. Cucunuba, R. Fitzjohn, K. A. M. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van-Elsand, E. Volz, H. Wang, Y. Wang, X. Xi, C. Donnelly, A. Ghani, N. Ferguson, Estimates of the severity of COVID-19 disease. medRxiv 2020.03.09.20033357 [preprint]. 13 March 2020.

19. Y. Liu, L.-M. Yan, L. Wan, T.-X. Xiang, A. Le, J.-M. Liu, M. Peiris, L. L. M. Poon, W. Zhang, Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect. Dis.* S1473-3099(20)30232-2 (2020). doi:10.1016/S1473-3099(20)30232-2 Medline

20. I. Dorigatti, L. Okell, A. Cori, N. Imai, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, R. FitzJohn, H. Fu, K. Gaythorpe, A. Hamlet, W. Hinsley, N. Hong, M. Kwun, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, R. Wang, C. Walters, X. Xi, C. Donnelly, A. Ghani, N. Ferguson, *Report 4: Severity of 2019-Novel Coronavirus (nCoV)* (10 February 2020); www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-severity-10-02-2020.pdf.

21. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. Meredith, A. S. Azman, N. G. Reich, J. Lessler, The incubation period of 2019-nCoV from publicly reported confirmed cases: Estimation and application. medRxiv 2020.02.02.20020016 [preprint]. 4 February 2020.

22. H. Nishiura, N. M. Linton, A. R. Akhmetzhanov, Serial interval of novel coronavirus (2019-nCoV) infections. medRxiv 2020.02.03.20019497 [preprint]. 17 February 2020.

23. N. C. Grassly, C. Fraser, Mathematical models of infectious disease transmission. *Nat. Rev. Microbiol.* **6**, 477–487 (2008). doi:10.1038/nrmicro1845 Medline

24. BDI Pathogen Dynamics Group, Digital contact tracing for SARS-COV-2 (2020); https://bdi-pathogens.shinyapps.io/covid-19-transmission-routes/.

25. P. Mozur, R. Zhong, A. Krolik, “In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags.” *New York Times* (1 March 2020); www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html.

26. S. Chen, J. Yang, W. Yang, C. Wang, T. Bärnighausen, COVID-19 control in China during mass population movements at New Year. *Lancet* **395**, 764–766 (2020). doi:10.1016/S0140-6736(20)30421-9 Medline

27. Nuffield Council on Bioethics, *Research in Global Health Emergencies: Ethical Issues* (28 January 2020); www.nuffieldbioethics.org/publications/research-in-global-health-emergencies.

28. S. Zhao, Q. Lin, J. Ran, S. S. Musa, G. Yang, W. Wang, Y. Lou, D. Gao, L. Yang, D. He, M. H. Wang, 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. *Int. J. Infect. Dis.* **92**, 214–217 (2020). doi:10.1016/j.ijid.2020.01.050 Medline

29. T. Zhou, Q. Liu, Z. Yang, J. Liao, K. Yang, W. Bai, X. Lu, W. Zhang, Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCoV. *J. Evid. Based Med.* **13**, 3–7 (2020). doi:10.1111/jebm.12376 Medline

30. T. Ganyani, C. Kremer, D. Chen, A. Torneri, C. Faes, J. Wallinga, N. Hens, Estimating the generation interval for COVID-19 based on symptom onset data. medRxiv 2020.03.05.20031815 [preprint]. 8 March 2020.

31. X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li, G. M. Leung, Temporal dynamics in viral shedding and transmissibility of COVID-19. medRxiv 2020.03.15.20036707 [preprint]. 18 March 2020.

32. D. Klinkenberg, C. Fraser, H. Heesterbeek, The effectiveness of contact tracing in emerging epidemics. *PLOS ONE* **1**, e12 (2006). doi:10.1371/journal.pone.0000012 Medline

33. M. Salathé, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, J. H. Jones, A high-resolution human contact network for infectious disease transmission. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22020–22025 (2010). doi:10.1073/pnas.1009094108 Medline

34. E. Yoneki, J. Crowcroft, EpiMap: Towards quantifying contact networks for understanding epidemiology in developing countries. *Ad Hoc Netw.* **13**, 83–93 (2014). doi:10.1016/j.adhoc.2012.06.003

35. K. A. Nguyen, Z. Luo, C. Watkins, in *Progress in Location-Based Services 2014*, G. Gartner, H. Huang, Eds. (Springer, 2015), pp. 63–78.

36. G. M. Leung, P.-H. Chung, T. Tsang, W. Lim, S. K. K. Chan, P. Chau, C. A. Donnelly, A. C. Ghani, C. Fraser, S. Riley, N. M. Ferguson, R. M. Anderson, Y.-L. Law, T. Mok, T. Ng, A. Fu, P.-Y. Leung, J. S. M. Peiris, T.-H. Lam, A. J. Hedley, SARS-CoV antibody prevalence in all Hong Kong patient contacts. *Emerg. Infect. Dis.* **10**, 1653–1656 (2004). doi:10.3201/eid1009.040155 Medline

37. K.-Q. Kam, C. F. Yung, L. Cui, R. Tzer Pin Lin, T. M. Mak, M. Maiwald, J. Li, C. Y. Chong, K. Nadua, N. W. H. Tan, K. C. Thoon, A Well Infant with Coronavirus Disease 2019 with High Viral Load. *Clin. Infect. Dis.* ciaa201 (2020). doi:10.1093/cid/ciaa201 Medline

38. X. Lu, L. Zhang, H. Du, J. Zhang, Y. Y. Li, J. Qu, W. Zhang, Y. Wang, S. Bao, Y. Li, C. Wu, H. Liu, D. Liu, J. Shao, X. Peng, Y. Yang, Z. Liu, Y. Xiang, F. Zhang, R. M. Silva, K. E. Pinkerton, K. Shen, H. Xiao, S. Xu, G. W. K. Wong, SARS-CoV-2 Infection in Children. *N. Engl. J. Med.* NEJMc2005073 (2020). doi:10.1056/NEJMc2005073 Medline

39. G. Kampf, D. Todt, S. Pfaender, E. Steinmann, Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *J. Hosp. Infect.* **104**, 246–251 (2020). doi:10.1016/j.jhin.2020.01.022 Medline

40. The code used for our analyses available at https://doi.org/10.5281/zenodo.3727255.

41. L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, D. G. Bonsall, C. Fraser, Quantifying dynamics of SARS-CoV-2 transmission suggests that epidemic control and avoidance is feasible through instantaneous digital contact tracing. medRxiv 2020.03.08.20032946 [preprint]. 12 March 2020.

42. China Centre for Disease Control, Distribution of New Coronavirus Pneumonia; http://2019ncov.chinacdc.cn/2019-nCoV/.

43. L. T. Phan, T. V. Nguyen, Q. C. Luong, T. V. Nguyen, H. T. Nguyen, H. Q. Le, T. T. Nguyen, T. M. Cao, Q. D. Pham, Importation and Human-to-Human Transmission of a Novel Coronavirus in Vietnam. *N. Engl. J. Med.* **382**, 872–874 (2020). doi:10.1056/NEJMc2001272 Medline

44. M. Ki, Task Force for 2019-nCoV, Epidemiologic characteristics of early cases with 2019 novel coronavirus (2019-nCoV) disease in Republic of Korea. *Epidemiol. Health* **42**, e2020007 (2020). doi:10.4178/epih.e2020007

45. C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, M. Seilmaier, C. Drosten, P. Vollmar, K. Zwirglmaier, S. Zange, R. Wölfel, M. Hoelscher, Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. *N. Engl. J. Med.* **382**, 970–971 (2020). doi:10.1056/NEJMc2001468 Medline