DECLINE IN GLOBAL TRANSMISSION RATES OF COVID-19

A PREPRINT

Ethan Romero-Severson
Theoretical biology and biophysics
Los Alamos National Laboratory
Los Alamos, NM
eoromero@lanl.gov

Nick Hengartner
Theoretical biology and biophysics
Los Alamos National Laboratory
Los Alamos, NM
nickh@lanl.gov

Grant Meadors
Computational physics
Los Alamos National Laboratory
Los Alamos, NM
gdmeadors@lanl.gov

Ruian Ke
Theoretical biology and biophysics
Los Alamos National Laboratory
Los Alamos, NM
rke@lanl.gov

April 23, 2020

ABSTRACT

Our method uses a hybrid deterministic and stochastic formalism that allows for time variable transmission rates and detection probabilities. The model was fit using iterated particle filtering to case count and death count time series from 51 countries. We found evidence for a declining transmission rate in 42 of the 51 examined countries. Of those 42 countries 34 have significant evidence for subcritical transmission rates, although the decline in new cases are relatively slow compared to the initial growth rates. This suggests that global scale social distancing efforts to slow the spread of COVID-19 are effective although they need to be strengthened in many regions and maintained in others to avoid further resurgence of COVID-19. The slow decline also suggests alternative strategies to control the virus are needed before social distancing efforts are partially relaxed.

***K*eywords** COVID-19 · Partially observed Markov process · Iterated particle filtering

1 Introduction

Since its initial outbreak in Wuhan China in late 2019 and early 2020 (1), COVID-19 caused repeated rapid outbreaks across the global from February through April 2020. The extremely rapid spread of COVID-19 in China (2) does not appear to be an anomaly: the disease has shown a short doubling time (2.4-3.6 days) outside of China as well (3). As of April 17, 2020, the virus caused 2,181,508 reported infections, and 147,337 deaths globally (4). In response, most affected countries/regions have implemented strong social distancing efforts, such as school closures, working-from-home, shelter-in-place orders. As a result, the spread of COVID-19 slows down substantially in some countries (5), and thus the epidemic curves are flattened. As social distancing induces high costs to both society and individuals, plans to relax social distancing are discussed. However, it is not clear if the transmission in these areas are subcritical and how long the transmission is reduced to a level that are manageable for health care systems.

In this report, we developed a deterministic-stochastic hybrid model and fitted the model to case incidence and death incidence time series data from 51 countries. Following the approach suggested by King et al.(6), we use a (partially) stochastic model and base our estimates on incidence rather than cumulative incidence data . Using both case count and death count of data allowed us to disentangle changes in surveillance intensity from changes in transmission (3). We found evidence for large decreases in the country-level transmission rates, in several of the worst-affected countries. Importantly, using data up to April 16, 2020, we specifically tested whether the transmission is subcritical (i.e. the reproductive number R is below 1). Most countries showed large decreases in transmission rates over time, and some

have transmission rates below the epidemic threshold, i.e. $R = 1$. On the other hand, many countries still appear to be showing rapid exponential growth. Given its highly contagious nature, COVID-19 can spread rapidly when strong social distancing measures are lifted, even partially (3). Alternative strategies that can effective control the virus are needed when social distancing measuares

2 Materials and Methods

2.1 Data

Case count and death data were downloaded from the Johns Hopkins GitHub repository (`https://github.com/CSSEGISandData/COVID-19`) on April 16 2020. Any country in the data that had more than 2000 cumulative cases by April 16 was included in the modeled data. China and South Korea were excluded from the final analysis because the model was not capable of capturing the complex dynamics in those countries. To minimize the effect of repatriated cases we started each time series on the first day when the cumulative number of cases was greater than or equal to 100. All data processing and model fitting were otherwise done on the incidence scale. To address obvious bulk-reporting issues in the data (e.g. sudden zeros in the data followed by very large numbers), we smoothed the data using Tukey's 3-median method. Because several countries had days with a single death surrounded by days with no deaths, which the smoothing method would set to zero, days with a single death were not replaced with smoothed values. The original data and the smoothed data used for estimation are shown in Supplemental figure S1.

2.2 Model

We model the spread of COVID-19 as a partially observed Markov process with real-valued states S (susceptible), E (exposed), I (infected), and R (recovered) to model the latent population dynamics, and integer-valued states C_0 (to be counted), Y_1 (counted cases), $D_{0:3}$ (dying), and Y_2 (counted deaths) to model sampling into the data. The model and all of its parameters have time units of days. The latent population model is governed by the following ordinary differential equations

$$
\begin{aligned}
\frac{dS}{dt} &= -\chi(t) I \frac{S}{N} \\
\frac{dE}{dt} &= \chi(t) I \frac{S}{N} - \lambda_{EI} E \\
\frac{dI}{dt} &= \lambda_{EI} E - \lambda_{IR} E \\
\frac{dR}{dt} &= \lambda_{IR} E
\end{aligned}
$$

where $\chi(t)$ is the time-variable transmission rate and $N = S + E + I + R$ is assumed to be fixed over the run of the model. λ_{EI} is the rate at which exposed persons become infectious and λ_{IR} is the rate at which cases recover (i.e. are either no longer infectious or die). At every time interval, we sample persons moving from the E to I states into a stochastic arm of the model that is used to calculate the likelihood of the data.

To relate the latent population model to data, we randomly sample individuals from the unobserved population into a stochastic process that models the random movement from infection to being either counted as an observed case or counted as an observed death. The number of persons sampled into the observation arm of the model over a time interval dt is given by $X \sim \text{Multinomial}\big(g(E\lambda_{EI}dt), \{(\rho^c(t)\omega^c, \rho^c(t)\omega, \rho(t)\omega^c, \rho(t)\omega\}\big)$ where $\rho(t)$ and ω are the probabilities that an infected person will be counted or die respectively, $\rho^c(t)$ and ω^c are the probabilities of being not sampled or not dying respectively, and $g(u)$ is a stochastic function that maps from real-value u to integer $g(u)$; it takes value $\lfloor u \rfloor$ with probability $u \mod 1$ and value $\lceil u \rceil$ with probability $1 - (u \mod 1)$. In plain English, the model tracks the random fate of each newly infected person as they move from the exposed to infected state with respect to eventually being observed as a case and/or a death in data. At each time step of length dt, the change in state space is

given by

$$
\begin{aligned}
C_0(t+dt) &= C_0(t) + X_3 + X_4 - F_t \\
Y_1(t+dt) &= Y_1(t) + F_t \\
D_0(t+dt) &= D_0(t) + X_2 + X_4 - H_{1t} \\
D_1(t+dt) &= D_1(t) + H_{1t} - H_{2t} \\
D_2(t+dt) &= D_2(t) + H_{2t} - H_{3t} \\
D_3(t+dt) &= D_3(t) + H_{3t} - H_{4t} \\
Y_2(t+dt) &= Y_2(t) + H_{4t}
\end{aligned}
$$

where $F_t|C_0$ is a random variate from $\text{Bin}(C_0, 1 - \exp(-\lambda_{Y_1} dt)$ and H_{it} is a random variate from $\text{Bin}\left(D_{i-1}, 1 - \exp\left(-\frac{4}{21}dt\right)\right)$. X_i indicates the ith element of Multinomial random variate defined above. The rate λ_{Y1} determines the rate at which persons who will be counted are counted (i.e. lower values of λ_{Y1} mean a longer delay in cases showing up in the data). The values of $Y_{1:2}$ are set to zero at the beginning of every day such that they accumulate the simulated number of cases and deaths that occur each day. Both the transmission rate and detection probability are allowed to vary with time in the following way

$$
\chi(t) = \begin{cases} \chi_0 & \text{if } t \leq t_\chi \\ \chi_0 + \frac{\chi_f - \chi_0}{t_f - t_\chi}(t_\chi - t) & \text{if } t_\chi < t \leq (t_f - 10) \\ \chi_f & \text{if } t > (t_f - 10) \end{cases}
$$

$$
\rho(t) = \begin{cases} \rho_0 & \text{if } t \leq t_\rho \\ \rho_0 + \frac{\rho_f - \rho_0}{t_f - t_\rho}(t_\rho - t) & \text{if } t_\rho < t \leq (t_f - 10) \\ \rho_f & \text{if } t > (t_f - 10) \end{cases}
$$

where t_f is the final time in the given dataset. Note that, to prevent over-fitting, the final values the transmission rate and detection probability are forced to be constant over the last 10 days of the data.

2.3 Parameter estimation

We assume that the data are Negative Binomial-distributed, conditional on the simulated number of cases and deaths that occur in a given time interval, i.e. the number of cases in the ith observation period has density $\text{NegBin}(Y_1, \epsilon_1)$ and the number of deaths in the ith observation period has density $\text{NegBin}(Y_2, \epsilon_2)$. The Negative Binomial is parameterized such that the first argument is the expectation and the second is an inverse overdispersion parameter that controls the variance of the data about the expectation; as ϵ_i becomes large, the data model approaches a Poisson with parameter Y_i. Both ϵ_1 and ϵ_2 were estimated from the data.

The model was fit to the data using an iterated particle-filtering method, implemented in the `pomp` R library (7). To optimize the likelihood of the data, we used 2000 particles in 100 iterations for each country. The reported likelihoods were measured using 6000 particles at the optimized Maximum Likelihood Estimates (MLEs). The Monte Carlo error in the likelihood measurement was less than 0.1 log units. Because we are using an optimization rather than sampling method we fit a sequence of models to test if the data support time-variable detection and/or transmission rates. The first model assumes both the transmission rate and detection probability are constant (i.e. $\chi_0 = \chi_f$ and $\rho_0 = \rho_f$); the second model allows for time variable detection probability (i.e. $\chi_0 = \chi_f$); the third model allows for time variable transmission rates $\rho_0 = \rho_f$; and the fourth model estimates all the free parameters. We determine the best model by a sequence of likelihood ratio tests. If model 2 and/or model 3 are supported at $\alpha = 0.05$, then the model with the higher likelihood is preferred. If there is support for model 4 over the preferred model, then model 4 is considered best. This testing cascade avoids overfitting by considering both possible explanations in isolation and only including both if the data support it. Finally, for data sets that support a declining transmission rate, we run an additional model where the final transmission rate is set such that the reproduction number is 1. Note that the model structure assumes the final transmission rate is constant over the last 10 days of the data. We use the likelihood from that model conduct a likelihood ratio test to distinguish between countries that have nearly subcritical transmission rates to ones that have clear evidence of a shrinking epidemic.

The initial state at time zero is computed by assuming there were $I(0)$ infected persons 21 days before the first reported death (by definition time one). The number of initial number of susceptibilities was assumed to be the predicted 2020 population size of the given country (8). The model is then simulated forward for 21 days, which is taken at the initial state of the model at time zero. Parameters have constraints $I(0) \in [1, 1000]$, $\rho_0 \in [0, 1]$, $\rho_f \in [0, 1]$, $\lambda_{Y1}^{-1} \in [5, 21]$ t_χ and $t_\rho \in [0, t_f - 10]$; all other parameters are constrained to be positive.

3 Results

We fit our model to data collected from 51 countries. Model fits are shown in Figure 1, and parameter values are given in Table 5. The model can capture the data well, with a few exceptions. Iran, Spain, and Italy all display a tight temporal correlation between changes in the number of cases and changes in the numbers of deaths (i.e. a downturn in cases co-occurring with a downturn in deaths), which the model cannot reproduce exactly as it occurs. This may due to the assumption in the model that the delay in case detection from infection is constant over time, and the inability of the model to find a single time delay in case detection that fits both the case and death time series in Italy, Spain, and Iran both at the beginning and end of the data suggests heterogeneity in the detection rate over time of either cases or deaths in these countries. Modeling efforts that focus on these regions should consider including temporal heterogeneity in the detection rate.

We found evidence for decreasing transmission rates in 42 of the 51 included countries. Figure 2 shows the starting and final exponential growth rates (growth rates were computed from model parameters according to (9)). Note that the starting exponential growth rate is the growth rate estimated from the initial data points we collected in each country, and it is not necessarily the rate of outbreak growth at the beginning of the outbreak. In countries with a decline, the average reduction in the transmission rate was 84% from the starting value. Of the included countries, 32 had shrinking epidemics (i.e. infected persons are generating less than 1 new infection on average). To further test whether there exists significant evidence that the transmission is sub-critical, we modified our model to force the final transmission to be at the threshold, i.e. the growth rate is 0, during the last 10 days of simulation. Comparing the two models, we found that 8 countries have maximum likelihoods that suggest $R < 1$ (i.e. negative growth) could not be meaningfully distinguished from $R = 1$. Countries with significantly subcritical dynamics on April 16 2020 were Algeria, Australia, Austria, Belgium, Czechia, Ecuador, France, Germany, Greece, Iran, Israel, Italy, Luxembourg, Malaysia, Netherlands, Norway, Poland, Portugal, Romania, Spain, Switzerland, Thailand, United Kingdom, and the United States. One caveat in the interpretation of the results is that the potentially long delays (the model allows up to a 21 day average delay) in reporting lead to a high level of ambiguity in the data concerning the transmission rate in the recent past.

We fixed the case fatality ratio to 1% ($\omega = 0.01$) because that parameter is not estimable from the data. To check the influence of that assumption we reoptimized the best model for each country assuming $\omega = 0.02$ or $\omega = 0.002$. The qualitative results were robust to this range of case fatality ratios. Countries that were found to have shrinking epidemics at $\omega = 0.01$ were also found to have shrinking epidemics at $\omega = 0.02$ and $\omega = 0.002$. Lower levels of the case fatality ratio let to higher fractions of the population being infected, but even at $\omega = 0.002$ this was not sufficient to explain the data through exhaustion of susceptibles alone.

The model estimated the time delay in the time from infection to detection (Supplemental Figure S4); the average delay was 14 days. We further looked for correlations in the initial detection probability and initial growth rate (Supplemental Figure S3), but the correlation appeared to be insignificant (removing two outliers, Qatar and South Africa, led to a non-significant negative correlation (p=0.32).

Our model included an explicit susceptible population in each country, allowed us to estimate the fraction of the population that had been infected by April 16 2020 (Supplemental Figure S5). On average 1.5% of the included countries infected population was infected on April 16 2020. Our optimization-based inference method does not allow us to easily compute confidence intervals on the parameter point estimates. Our previous work using a deterministic modeling approach suggests the range of variation could be an order of magnitude given uncertainties in parameter values, such as the infection case fatality ratio (3).

4 Discussion

Our model found evidence for reductions in transmission rates of COVID-19 in 42 of 51 examined countries. Encouragingly, of those countries, we found statistical evidence that the transmission of COVID-19 is decreasing in 24 countries, i.e. the effective reproductive number is less than 1, using data up to April, 16, 2020. This suggests that, despite the highly heterogeneous populations represented by these countries, the growth of COVID-19 outbreak can be reverted. Although our model cannot attribute exact causes to the global decline in transmissions rates, most countries implemented sustained, population-level social distancing efforts over a period of weeks to months. These efforts are highly likely to play a major role in reducing the transmission of COVID-19 (5).

We estimated that in countries with decreasing transmission, the rate of decrease is in general less than 0.1/day. Based on data from 8 European countries, the US and China, we previously estimated that in the absence of intervention efforts, the epidemic can grow at rates between 0.19-0.29/day (2). This means that the outbreak can grow rapidly and avert public heath efforts made if social distancing measures are completely relaxed. For example, if the rate of decrease under strong public health interventions is 0.1/day and the growth rate in the absence of public heath interventions

is 0.2/day, then, the number of cases averted in two weeks of intervention will be regained in only one week. Social distancing measures have their own social costs. Our results suggests alternative strategies to control the virus are needed in place when social distancing efforts are relaxed. Due to the uncertainties in the impact of each specific control measures, changes to policies should be made slowly because the signal of changing transmission can take weeks to fully propagate into current data streams as a result of the long lag between infection to case confirmation (as we estimated to be approximately 2 weeks).

Overall, our results suggest that COVID-19 is controllable in diverse settings using a full range of strong and comprehensive non-pharmaceutical measures, and that future deaths from the disease are avoidable.

5 Tables

Symbol	Description	Value
λ_{EI}	Latency rate	$\frac{1}{3}$
λ_{IR}	Recovery rate	$\frac{1}{10}$
ω	Case fatality ratio	1%
χ_0	Initial transmission rate	fit
t_χ	Time of trans. rate change	fit
χ_f	Final transmission rate	fit
ρ_0	Initial probability of being counted	fit
t_ρ	Time of counting prob. change	fit
ρ_f	Final counting prob.	fit
$I(0)$	Initial number infected	fit

Country	Model	I_0	χ_0	χ_f	p,$\chi_f = 0$	r_0	r_f	ρ_0	ρ_f	λ_{Y1}^{-1}
Algeria	3	731	0.34	0.01	0.02	0.14	-0.09	0.05	0.05	18
Argentina	3	952	0.22	0.05	0.69	0.08	-0.04	0.15	0.15	13
Australia	3	19	0.44	<0.01	<0.01	0.18	-0.1	0.41	0.41	5
Austria	4	14	0.57	<0.01	<0.01	0.23	-0.1	0.35	<0.01	9
Belgium	4	14	0.69	0.01	<0.01	0.28	-0.08	0.12	0.03	21
Brazil	3	21	0.7	0.03	1	0.28	-0.06	0.1	0.1	18
Canada	3	8	0.68	0.04	0.39	0.27	-0.05	0.18	0.18	13
Chile	3	26	0.49	0.04	0.15	0.2	-0.05	0.46	0.46	15
Colombia	3	501	0.26	<0.01	0.1	0.1	-0.1	0.12	0.12	9
Czechia	3	47	0.45	<0.01	<0.01	0.19	-0.1	0.2	0.2	6
Denmark	2	987	0.22	0.22	<0.01	0.08	0.08	0.19	<0.01	16
Dominican Republic	3	997	0.26	0.1	1	0.1	<0.01	0.15	0.15	18
Ecuador	3	870	0.32	<0.01	0.01	0.13	-0.1	0.14	0.14	18
Egypt	3	896	0.22	0.14	<0.01	0.08	0.03	0.05	0.05	5
Finland	2	569	0.18	0.18	<0.01	0.06	0.06	0.41	0.05	19
France	3	51	0.64	<0.01	0.01	0.26	-0.1	0.06	0.06	13
Germany	3	5	0.73	<0.01	<0.01	0.29	-0.1	0.21	0.21	7
Greece	3	954	0.18	0.01	<0.01	0.05	-0.09	0.18	0.18	8
India	3	85	0.41	0.21	0.22	0.17	0.08	0.06	0.06	5
Indonesia	4	969	0.27	0.24	<0.01	0.1	0.09	0.05	0.03	5
Iran	3	852	0.39	<0.01	<0.01	0.16	-0.1	0.13	0.13	18
Ireland	2	217	0.35	0.35	<0.01	0.14	0.14	0.12	0.03	5
Israel	3	3	0.61	0.02	<0.01	0.25	-0.08	0.49	0.49	9
Italy	4	282	0.55	<0.01	<0.01	0.23	-0.1	0.08	0.02	21
Japan	3	457	0.16	0.32	<0.01	0.04	0.13	0.18	0.18	5
Luxembourg	3	762	0.2	<0.01	<0.01	0.07	-0.1	0.41	0.41	5
Malaysia	3	58	0.34	0.02	<0.01	0.14	-0.07	0.48	0.48	12
Mexico	3	682	0.31	0.15	<0.01	0.12	0.03	0.08	0.08	19
Netherlands	3	154	0.52	<0.01	<0.01	0.21	-0.1	0.08	0.08	21
Norway	3	125	0.32	0.02	<0.01	0.13	-0.08	0.3	0.3	5
Pakistan	3	738	0.2	0.04	0.17	0.07	-0.05	0.43	0.43	16
Panama	3	834	0.22	<0.01	0.2	0.08	-0.1	0.27	0.27	21
Peru	1	289	0.31	0.31	<0.01	0.13	0.13	0.13	0.13	15
Philippines	3	403	0.31	0.11	1	0.12	0.01	0.06	0.06	5
Poland	3	499	0.28	<0.01	0.01	0.11	-0.1	0.2	0.2	20
Portugal	3	89	0.5	<0.01	<0.01	0.21	-0.1	0.24	0.24	21
Qatar	3	923	0.06	0.38	<0.01	-0.03	0.16	0.8	0.8	6
Romania	3	20	0.59	0.03	0.01	0.24	-0.06	0.13	0.13	20
Russia	3	25	0.47	0.32	<0.01	0.2	0.13	0.41	0.41	16
Saudi Arabia	1	289	0.24	0.24	<0.01	0.09	0.09	0.22	0.22	7
Serbia	2	237	0.29	0.29	<0.01	0.12	0.12	0.26	0.12	19
Singapore	3	8	0.26	0.24	<0.01	0.1	0.09	0.69	0.69	5
South Africa	2	559	0.15	0.15	<0.01	0.03	0.03	0.89	0.41	7
Spain	3	172	0.6	0.01	<0.01	0.25	-0.08	0.03	0.03	7
Sweden	4	624	0.28	<0.01	1	0.11	-0.1	0.17	0.06	20
Switzerland	4	42	0.53	<0.01	<0.01	0.22	-0.1	0.2	<0.01	11
Thailand	4	14	0.45	<0.01	0.02	0.19	-0.1	0.59	<0.01	9
Turkey	3	441	0.49	0.12	0.36	0.2	0.02	0.22	0.22	20
United Arab Emirates	3	10	0.44	0.19	1	0.18	0.06	0.28	0.28	5
United Kingdom	3	20	0.76	0.05	0.03	0.3	-0.04	0.06	0.06	21
US	4	21	0.72	0.02	<0.01	0.29	-0.07	0.15	0.15	20

Table 1: Table of key parameter values. The 'Model' column referrers to the best model defined in the model selection procedure. The values r_0 and r_f are the initial and final exponential growth rates implied by the other model parameters. The growth rate r and transmission rate χ are derived according to (9). The term λ_{Y1}^{-1} is the mean time in days from infection to detection of a new case in days. χ_0, χ_f, r_0, r_f, have day^{-1} units; I_0 is a count; ρ_0 and ρ_f are probabilities; and λ_{Y1}^{-1} has day units.

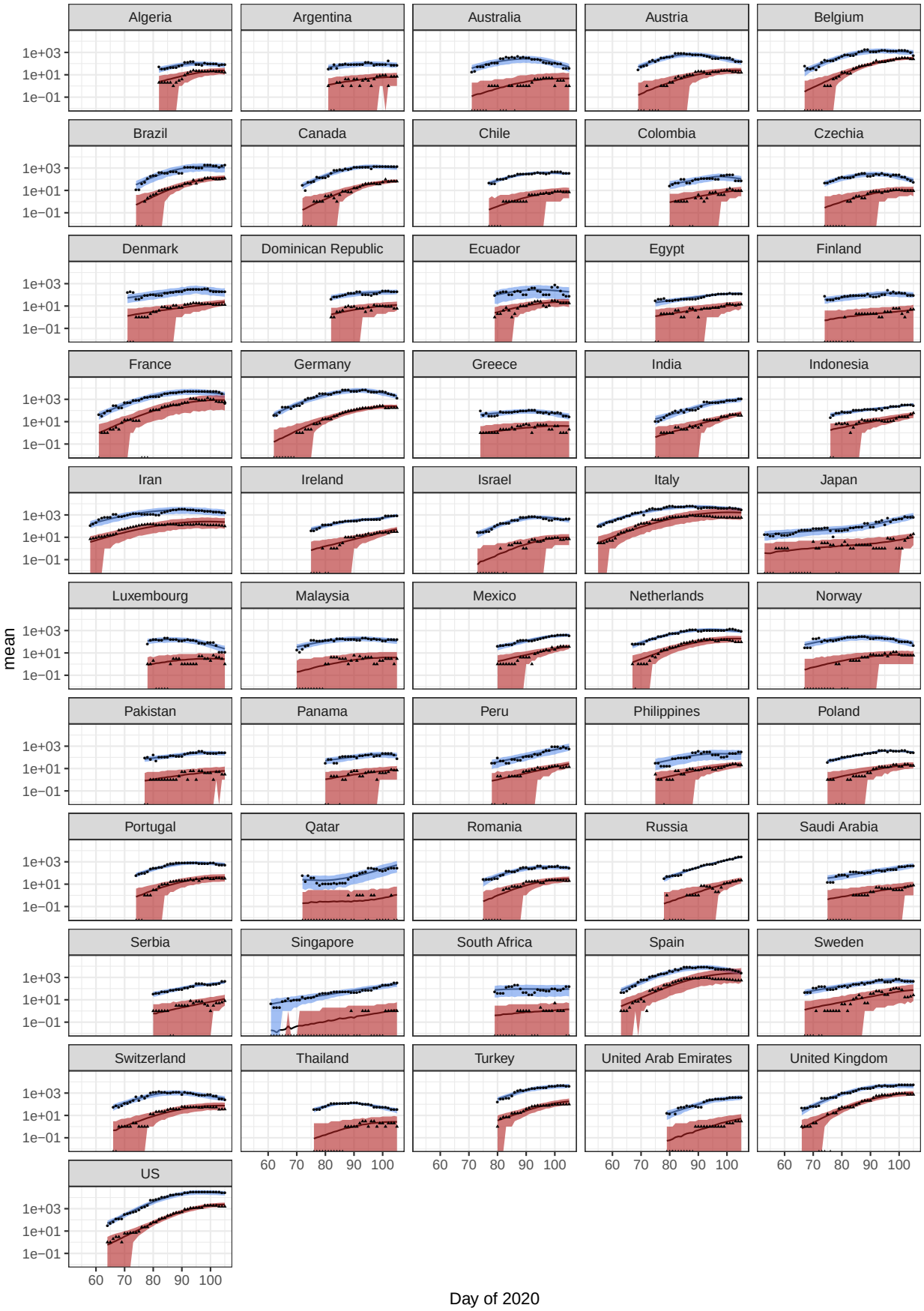

Figure 1: Model fits. Smoothed data points are shown in black and the expectations from the model are shown as solid lines. The envelope shows the 95% CIs for the data. The x-axis is measured in sequential days from the first datum.

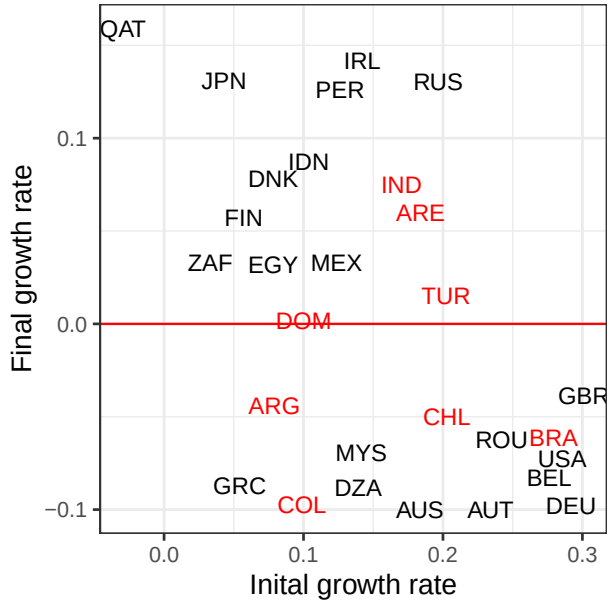

Figure 2: Values of the initial and final exponential growth rates as of April 16 2020. The red line delineates exponential growth from exponential decay. Countries below this line have a shrinking epidemic. The final growth rate of countries colored in red was found to be non-significantly different from zero. Country names are indicated by ISO Alpha-3 codes. Overlapping country names were omitted, figure with full data in Supplemental Figure S2.

6 Supplemental

6.1 Estimating the infection and counting rates

The differential equations for infected I and counted C cases in time t, with infection growth rate α and counting rate $\lambda > 0$, where both I and C are dimensionless numbers, but α and λ have units of $(\mathrm{time})^{-1}$, state that,

$$\frac{dI}{dt} = (I + C)\alpha - I\lambda, \tag{1}$$

$$\frac{dC}{dt} = I\lambda. \tag{2}$$

Equations 1 and 2 are coupled first-order equations. Cases where λ varies in time are of interest. As such cases are not necessarily integrable, a start is to first solve the constant λ case. Small perturbations can subsequently be probed. The equations can be rearranged into an uncoupled, second-order homogeneous ordinary differential equation, assuming constant α,

$$\frac{d^2 I}{dt} + (\lambda - \alpha)\frac{dI}{dt} + \left(\frac{d\lambda}{dt} - \alpha\lambda\right) I = 0, \tag{3}$$

which can be solved analytically for constant λ.

6.1.1 Constant counting rate

The solution is fixed by requiring that $I_0 \equiv I(t = 0)$ and $C_0 \equiv C(t = 0)$:

$$I(t) = \frac{1}{1 + \frac{\lambda}{\alpha}} \left([I_0 + C_0]\, e^{\alpha t} + \left[\frac{\lambda}{\alpha} I_0 - C_0\right] e^{-\lambda t} \right), \tag{4}$$

$$C(t) = \frac{1}{1 + \frac{\lambda}{\alpha}} \left(\frac{\lambda}{\alpha} [I_0 + C_0]\, e^{\alpha t} + \left[C_0 - \frac{\lambda}{\alpha} I_0\right] e^{-\lambda t} \right). \tag{5}$$

The full time derivatives of I and C satisfy Equations 1 and 2,

$$\frac{dI}{dt} = \frac{1}{1 + \frac{\lambda}{\alpha}} \left(\alpha [I_0 + C_0]\, e^{\alpha t} + \lambda \left[C_0 - \frac{\lambda}{\alpha} I_0\right] e^{-\lambda t} \right), \tag{6}$$

$$\frac{dC}{dt} = \frac{\lambda}{1 + \frac{\lambda}{\alpha}} \left([I_0 + C_0]\, e^{\alpha t} + \left[\frac{\lambda}{\alpha} I_0 - C_0\right] e^{-\lambda t} \right). \tag{7}$$

6.1.2 Long-duration limit

In the limit $t \to \infty$, these equations simplify to yield a ratio (I/C):

$$\lim_{t\to\infty} \left(\frac{dI}{dC}\right) = \lim_{t\to\infty} \left(\frac{dI/dt}{dC/dt}\right), \tag{8}$$

$$= \frac{\alpha}{\lambda},$$

as well as, where $\bar{I} \equiv \lim_{t\to\infty} I$ and $\bar{C} \equiv \lim_{t\to\infty} C$,

$$\bar{I} = \frac{I_0 + C_0}{1 + \frac{\lambda}{\alpha}} e^{\alpha t}, \tag{9}$$

$$= \left(\frac{\alpha}{\lambda}\right) \bar{C}.$$

This limit can also be interpreted as a time lag τ under the assumption of constant growth α, where $\tau = \alpha^{-1} \log(\alpha/\lambda)$. As expected of the growth rate, the limit also yields $d\bar{I}/dt = \alpha\bar{I}$ and $d\bar{C}/dt = \alpha\bar{C}$, the latter equivalent to $\lambda\bar{I}$.

6.1.3 Perturbation to counting rate

Taking this $t \to \infty$ limit, with $\lambda\bar{I} = \alpha\bar{C}$ assumed from Equation 9, discrepancies appear if λ is allowed to vary,

$$\lim_{t\to\infty}\left(\frac{dI}{dt}\right) = \alpha I - \frac{I}{\alpha\left(1+\frac{\lambda}{\alpha}\right)}\frac{d\lambda}{dt}, \tag{10}$$

$$\lim_{t\to\infty}\left(\frac{dC}{dt}\right) = \lambda I + \frac{I}{\alpha\left(1+\frac{\lambda}{\alpha}\right)}\frac{d\lambda}{dt}, \tag{11}$$

where the terms in $d\lambda/dt$ perturb the solutions to Equations 1 and 2. The perturbation to the differential equations is negligible when $d\lambda/dt \ll \alpha^2(1+\alpha^{-1}\lambda)$ and $\ll \alpha\lambda(1+\alpha^{-1}\lambda)$. If the testing rate λ is smaller than the growth rate α, the latter condition is stronger. When $\lambda \ll \alpha$, the condition simplifies to $d\lambda/dt \ll \alpha\lambda$.

Under the $(\alpha\lambda)$ condition, the natural logarithm of Equation 9 can be differentiated,

$$\log\bar{I} = \log\left(\frac{\alpha}{\lambda}\right) + \log\bar{C}, \tag{12}$$

$$\frac{d\log\bar{C}}{dt} - \frac{d\log\bar{I}}{dt} = \frac{1}{\lambda}\frac{d\lambda}{dt}, \tag{13}$$

where the $\alpha\lambda$ condition implies that the right-hand side is much less than α. As λ is positive, the time-derivative of the logarithm of $\bar{C}$ is greater than that of $\bar{I}$. The count $\bar{C}$ will increase faster than the number of uncounted, infected cases $\bar{I}$. This result holds when $d\log\bar{C}/dt \approx d\log\bar{I}/dt$ and both are approximately α. The interpretation is that time-varying λ is an acceptable perturbation if the difference between the time-derivatives of the logarithms of $\bar{C}$ and $\bar{I}$ is much smaller than the time-derivatives themselves.

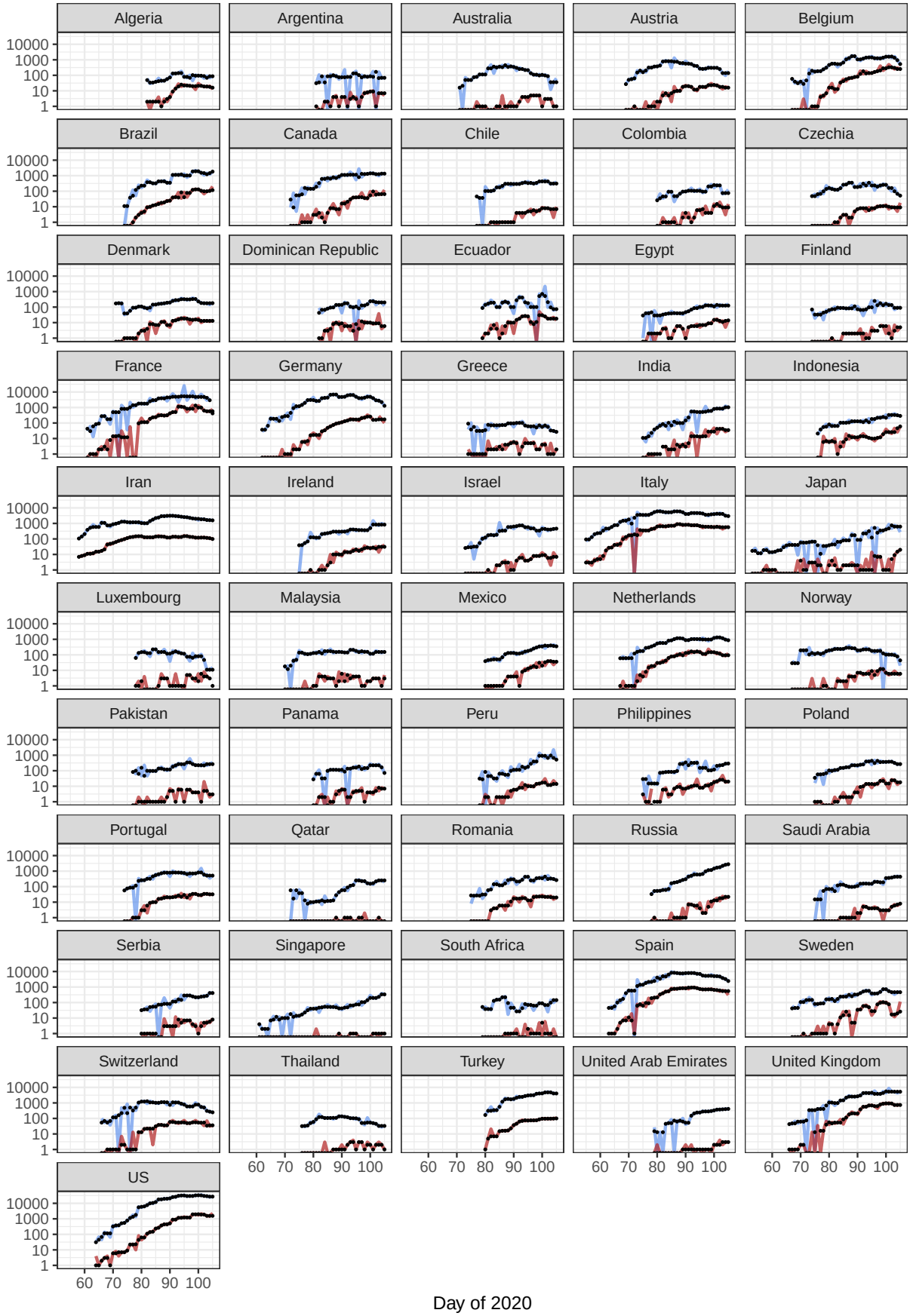

Figure S1: Raw and smoothed data. Raw data for cases (blue lines) and deaths (red lines) are shown with the smoothed time series (black) used for model fitting. The x-axis is measured in sequential days of 2020.

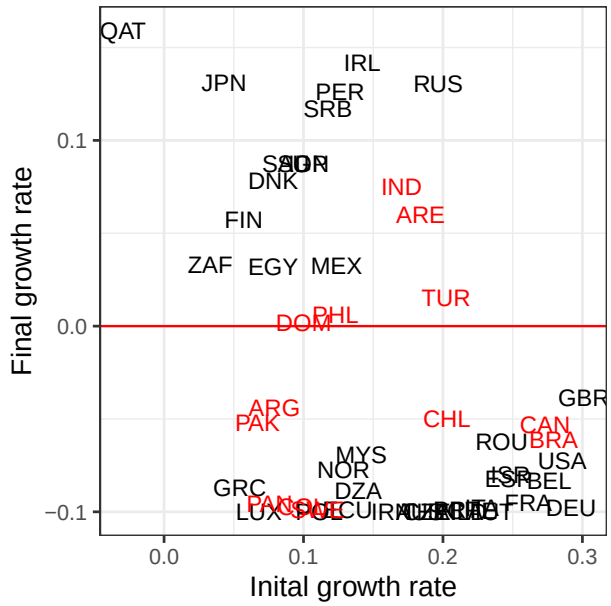

Figure S2: Values of the initial and final exponential growth rates as of April 16 2020. The red line delineates exponential growth from exponential decay. Countries below this line have a shrinking epidemic. The final growth rate of countries colored in red was found to be non-significantly different from zero. Country names are indicated by ISO Alpha-3 codes.

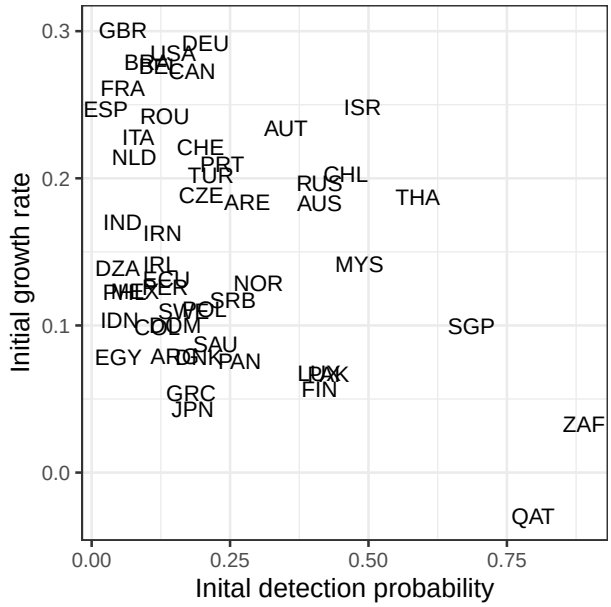

Figure S3: Values of initial detection probability and initial growth rate. Initial growth rate is negatively correlated (-0.35) with initial detection probability (p=0.01) Country names are indicated by ISO Alpha-3 codes.

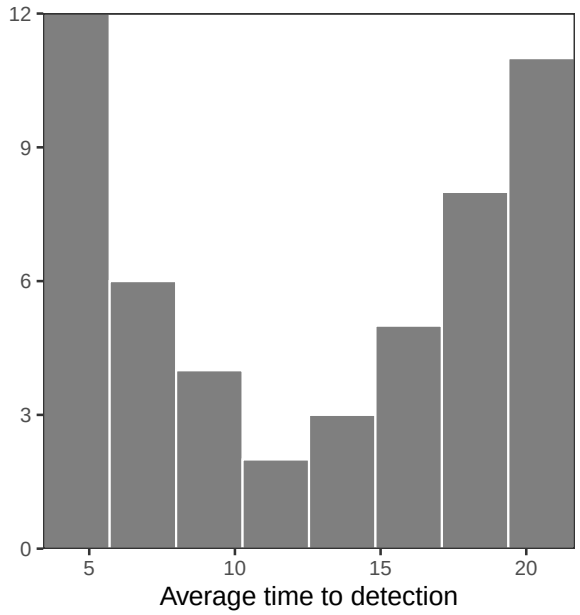

Figure S4: Histogram of the average time from infection to detection. Detection rates were bound such that the mean time to detection was between 5 and 21 days inclusive.

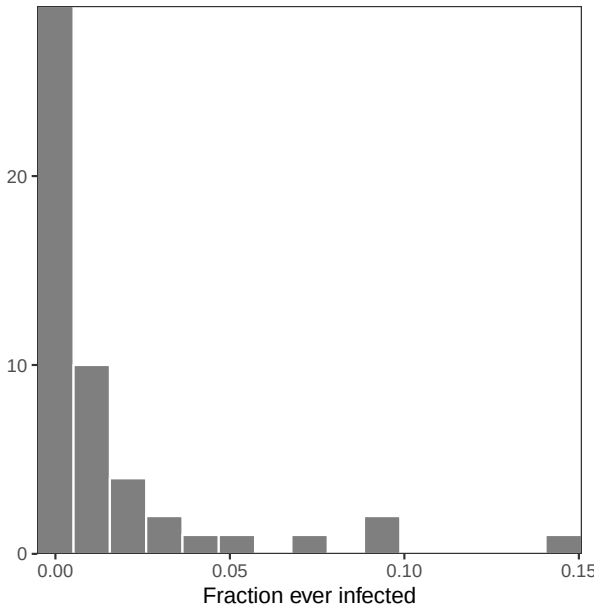

Figure S5: Fraction of population that has been infected by April 16 2020. Infected fraction includes exposed, infected, recovered, and dead.

References

[1] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. 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. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, and Z. Feng, "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia," *New England Journal of Medicine*, vol. 382, no. 13, pp. 1199–1207, Mar. 2020, publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2001316. [Online]. Available: https://doi.org/10.1056/NEJMoa2001316

[2] S. Sanche, Y. T. Lin, C. Xu, E. Romero-Severson, N. Hengartner, and R. Ke, "Early Release - High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2 - Volume 26, Number 7—July 2020 - Emerging Infectious Diseases journal - CDC." [Online]. Available: https://wwwnc.cdc.gov/eid/article/26/7/20-0282_article

[3] R. Ke, S. Sanche, E. Romero-Severson, and N. Hengartner, "Fast spread of COVID-19 in Europe and the US and its implications: even modest public health goals require comprehensive intervention," *medRxiv*, p. 2020.04.04.20050427, Apr. 2020, publisher: Cold Spring Harbor Laboratory Press. [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.04.04.20050427v1

[4] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet Infectious Diseases*, vol. 0, no. 0, Feb. 2020, publisher: Elsevier. [Online]. Available: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30120-1/abstract

[5] S. Flaxman, S. Mishra, and A. Gandy, "2020-03-30-COVID19-Report-13.pdf," Imperial College London, Tech. Rep., Mar. 2020. [Online]. Available: https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-30-COVID19-Report-13.pdf

[6] A. A. King, M. D. d. Cellès, F. M. G. Magpantay, and P. Rohani, "Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola," *Proc. R. Soc. B*, vol. 282, no. 1806, p. 20150347, May 2015. [Online]. Available: http://rspb.royalsocietypublishing.org/content/282/1806/20150347

[7] A. King, E. Ionides, C. Bretó, S. Ellner, B. Kendall, H. Wearing, M. Ferrari, M. Lavine, and D. Reuman, *pomp: Statistical inference for partially observed Markov processes (R package)*, 2009.

[8] C. I. A. Washington DC, "The World Factbook 2020." [Online]. Available: https://www.cia.gov/library/publications/resources/the-world-factbook/index.html

[9] J. Ma, "Estimating epidemic exponential growth rate and basic reproduction number," *Infectious Disease Modelling*, vol. 5, pp. 129–141, Jan. 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6962332/