State-level impact of social distancing and testing on COVID-19 in the United States

Weihsueh A. Chiu, Ph.D[1]✉, Rebecca Fischer, PhD, MPH[2], Martial L. Ndeffo-Mbah, Ph.D[1,2]✉

[1]Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77845

[2]Department of Epidemiology and Biostatistics, School of Public Health, Texas A&M University, College Station, TX 77845

✉To whom correspondence should be sent.

Telephone: +1 (979) 845 - 5646

Email: m.ndeffo@tamu.edu and wchiu@tamu.edu

Address: Texas A&M University
College of Veterinary Medicine & Biomedical Sciences
660 Raymond Stotzer Pkwy
College Station, TX 77845

Running Head: Social distancing, testing, and contact tracing to control COVID-19

Words counts: Abstract: 236, Main Text: 4291.

Keywords: COVID-19; social distancing; testing; contact tracing; mathematical modeling; Bayesian analysis

Ethical approval: Ethical approval was not required for this work.

Data sharing: No additional data available.

Competing interests: The authors declare no competing interests.

Authors' Contributions

WAC and MLNM conceived and designed the study and developed the model. WAC performed the analysis. WAC and MLNM drafted the paper, with critical revision by RF.

Abstract:

Social distancing measures have been implemented in the United States (US) since March 2020, to mitigate the spread of SARS-CoV-2, the causative agent of COVID-19. However, by mid-May most states began relaxing these measures to support the resumption of economic activity, even as disease incidence continued to increase in many states. To evaluate the impact of relaxing social distancing restrictions on COVID-19 dynamics and control in the US, we developed a transmission dynamic model and calibrated it to US state-level COVID-19 cases and deaths from March to June 20th, 2020, using Bayesian methods. We used this model to evaluate the impact of reopening, social distancing, testing, contact tracing, and case isolation on the COVID-19 epidemic in each state. We found that using stay-at-home orders, most states were able to curtail their COVID-19 epidemic curve by reducing and achieving an effective reproductive number below 1. But by June 20th, 2020, only 19 states and the District of Columbia were on track to curtail their epidemic curve with a 75% confidence, at current levels of reopening. Of the remaining 31 states, 24 may have to double their current testing and/or contact tracing rate to curtail their epidemic curve, and seven need to further restrict social contact by 25% in addition to doubling their testing and contact tracing rates. When social distancing restrictions are being eased, greater state-level testing and contact tracing capacity remains paramount for mitigating the risk of large-scale increases in cases and deaths.

The novel coronavirus pandemic (COVID-19) emerged in Wuhan, China in December 2019 and has now reached pandemic status, with spread to more than 210 countries and territories, including the United States (US) [1]. The US reported its first imported case of COVID-19 on January 20, 2020, arriving via an international flight from China [2]. Since then, the disease has spread rapidly within the US, with every state reporting confirmed cases within three weeks of the first reported community transmission. As of June 15th, the US has exceeded 2.1 million cases and 115,000 deaths, heterogeneously distributed across all states [1]. So far, states such as New York and New Jersey have borne the highest burden with more than 379,000 cases and 30,000 deaths and 166,000 cases and 12,000 deaths, respectively, while Montana and Alaska have each reported less than 700 cases and 20 deaths each [1].

COVID-19 is caused by a newly described and highly transmissible SARS-like coronavirus (SARS-CoV-2). Severe clinical outcomes have been observed with approximately 20% of symptomatic cases [3,4]. There is no vaccine and no cure or approved pharmaceutical intervention for this disease, making the fight against the pandemic reliant on non-pharmaceutical interventions (NPIs). These NPIs include: case-driven measures such as testing, contact tracing, and isolation [5]; personal preventive measures such as hand hygiene, cough etiquette, face mask use, eye protection, physical distancing, and surface cleaning, which aim to reduce the risk of transmission during contact with potentially-infectious individuals [6]; and social distancing measures to reduce interpersonal contact in the population. In the US, social distancing measures have included policies and guidelines to close schools and workplaces, cancel and restrict mass gatherings and group events, restrict travel, maintain physical separation from others (e.g. keeping six feet distance), and stay-at-home orders [7].

NPIs and other responses to COVID-19, especially stay-at-home orders, have varied widely across states, leading to spatial and temporal variation in the timing and implementation of mitigation strategies. This variation in policies and response efforts may have contributed to the observed heterogeneity in COVID-19 morbidity and mortality across states [8]. Recent studies suggest that statewide social distancing measures have likely contributed to reducing the spread COVID-19 epidemic in the US [9,10]. Understanding the extent to which NPIs, such as social distance, testing, contact tracing, and self-quarantine, influence COVID-19 transmission in a local context is pivotal for predicting the future course of the epidemic on a state-by-state basis. This in turn will inform how these NPIs should be optimized to mitigate the spread and burden of COVID-19 while awaiting development of pharmaceutical interventions (e.g. therapeutics and vaccines).

After several weeks of statewide stay-at-home orders, most US states have begun to ease their social distancing requirements [11], while attempting to increase their testing and contact tracing capacities [12]. Mathematical modeling is a unique tool to help answer these important and timely questions. Models can contribute valuable insight for public health decision-makers by providing an evaluation of the effectiveness of ongoing control strategies along with predictions of the potential impact of various policy scenarios [13].

To address these needs, we developed and validated a data-driven transmission dynamic model to evaluate the impact of social distancing, state-reopening, testing, and contact tracing on the state-level dynamics of COVID-19 infections and mortality in the US. We evaluated the transmissibility of COVID-19 in each state from March, 2020 to early June, 2020, to estimate the state-level impact of shelter-in-place and reopening on COVID-19 transmission. Finally, we evaluated the degree to which increasing testing efforts (rate of identification of infected cases) and/or contact tracing could curtail the spread of the diseases and enable greater relaxation of social distancing restrictions while preventing a resurgence of infections and deaths.

Results

Model performance and validation

We fit our model to state-level daily cases and deaths data using a Bayesian inference approach (see Online Methods). Model performance assessment for several representative states are shown in Figure 1, with full results in Figures S3 and S4. With respect to validation, the posterior 95% credible interval of our model projections, estimated using data through April 30th, 2020, covered 78% of the data points from May 1st through June 20th, 2020. Model performance for fitting all data through June 20th is shown in Figures S5-S7.

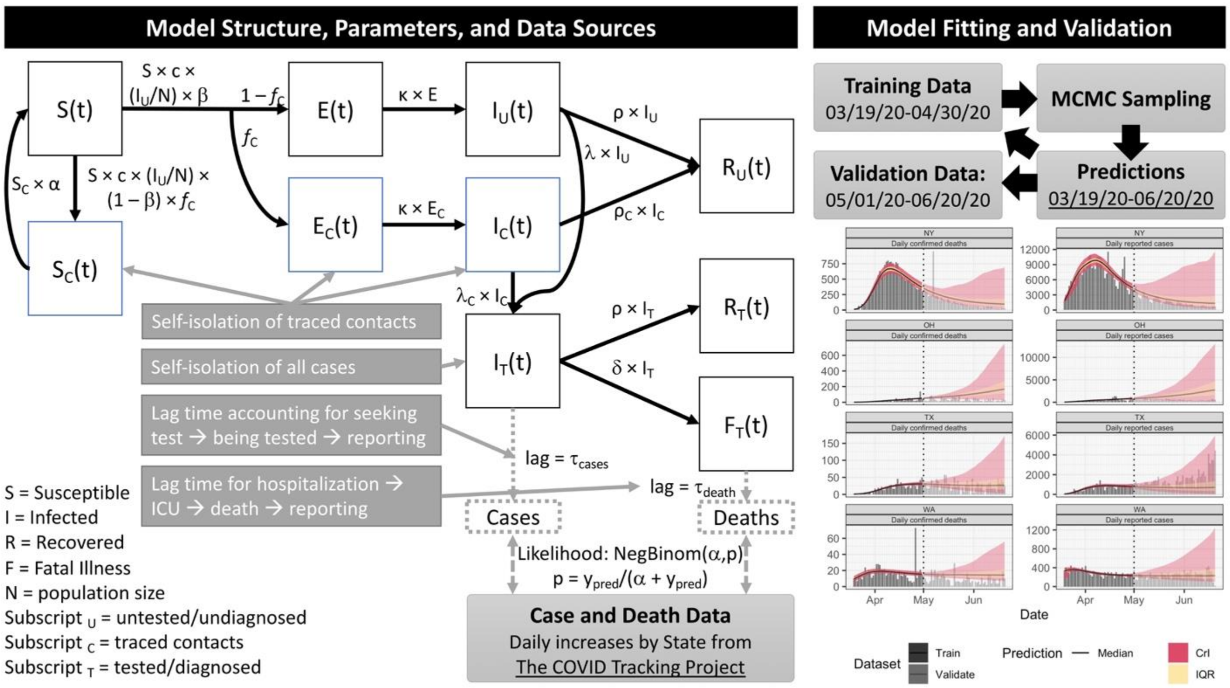

Figure 1. SEIR model structure, parameter, data sources, and fitting/validation methods. We fitted the model to daily reported cases and confirmed deaths from March 19th to April 30th and validated its projections against data from May 1st to June 20th. On the model projections, the black solid line is the median, the pink band is the 95% credible interval (CrI) and the orange is the inter-quartile range (IQR). We show model fitting and validation for four states: New York (NY), Ohio (OH), Texas (TX), and Washington (WA).

Estimations of effective reproduction number

Using the posterior distribution of our model parameters we estimated the effective reproduction number R_{eff} from March 19th to June 20th, 2020 and identified the minimum level of transmission achieved in each state (Figure 2A). We found that for all, except two states (Arkansas and Utah), the minimum R_{eff} value was less than 1 and these values were mainly achieved during the state shelter-in-place (Figure 2A). On June 20th, 2020, 27 states had at least a 0.5 probability that R_{eff}>1. Thus, the model predicts that as states are reopening, a majority of states are at risk of continued increases in the scale of the outbreak and require additional mitigation to contain the spread of the disease.

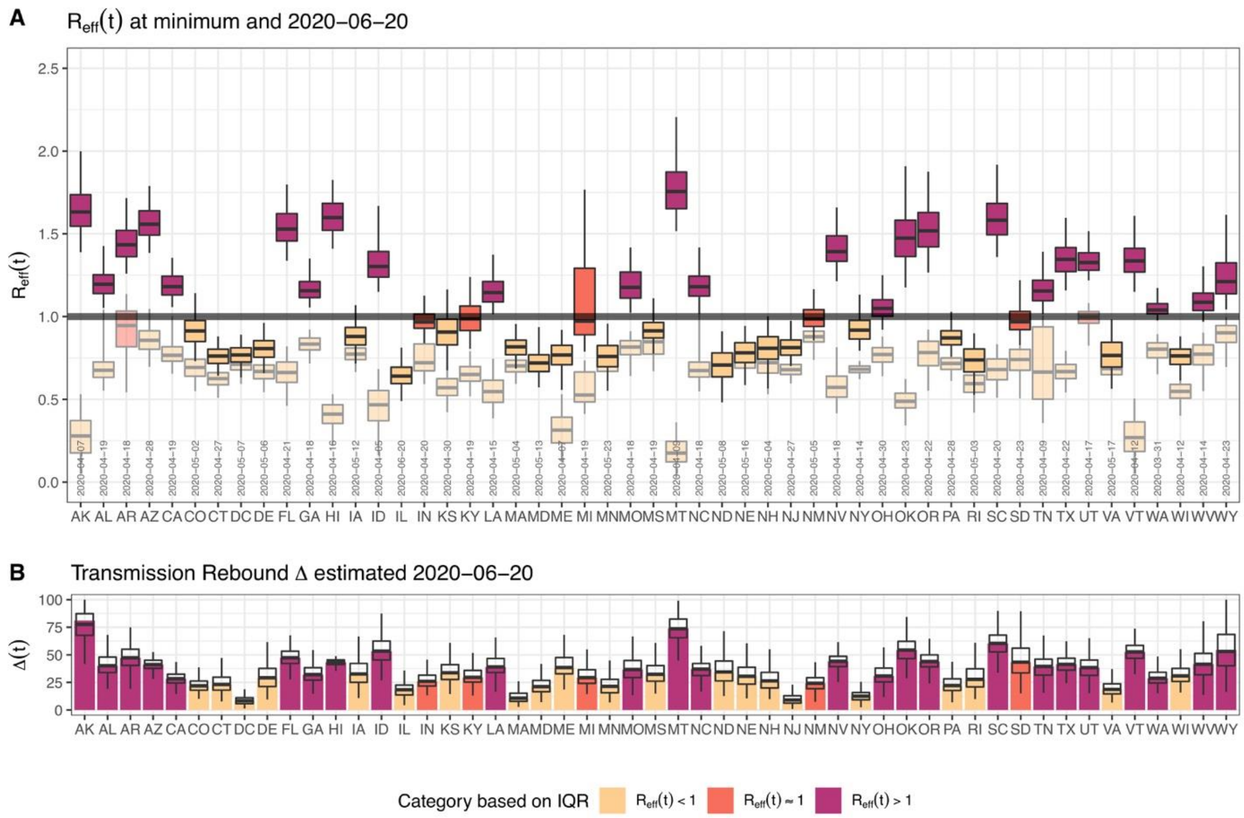

Figure 2. Estimated effective reproduction number R_{eff} and the level of reopening/rebound in transmission as of June 20th, 2020 for all states. (A) shows estimated R_{eff} (median, IQR, and 95% CrI) across *States. The figure shows that "now" (value on June 20th, 2020) and the "minimum" (between March 19th, 2020 and June 20th, 2020) in lighter shades of each color. It also includes the date of the minimum R_{eff}. (B) shows the level of reopening/rebound in disease transmission in each state relative to its minimum value during state shelter-in-place (median, IQR, and 95% CrI).*

We conducted an analysis of variance to evaluate the contribution of each parameter to the variation in R_{eff} value (Table S1). Across states, we found that the largest drivers of variation in R_{eff} are the power parameter for social distancing, η, the maximum relative increase in contact after shelter-in-place orders, r_{max}, and the fraction of contact traced, f_C, which together contribute over 60% of variance (Figure S8). This observation is consistent with mobility data alone being insufficient to account for the combined effect of multiple control measures, and suggest that the degree of adoption of non-mobility-related measures, such as enhanced hygiene practices and contact tracing, play a large role in the extent to which a state may reduce disease transmission.

For each state, we also estimated the current level of reopening/rebound Δ in disease transmission relative to its lowest transmission rate observed during shelter-in-place (Figure 2B). We found that only nine states had a 50% or more rebound in COVID-19 transmission by June 20th, 2020 while eight states had a 25% or less rebound in transmission (Figure 2B).

Impact of testing and contact tracing on easing of social distancing

Bringing and keeping the effective reproduction number, R_{eff}, below 1 is necessary and sufficient to curtail the spread of an outbreak. We evaluated the probability of keeping R_{eff}<1 for different levels of testing and contact tracing under the June 20th, 2020 level of state reopening. We found that for 12 states and the District of Columbia have at least 0.975 probability of keeping R_{eff}<1, and 22 states have less than 0.025 probability of bringing and keeping R_{eff}<1, under their current level of testing and contact tracing (Figure S9). We found that for most states bringing and keeping R_{eff}<1 may not be possible without increase contact tracing efforts as increasing testing and isolation alone would be sufficient or require extremely high coverage to curtail the epidemic curve with a 0.975 probability (Figure S9).

To evaluate the impact of scaling up testing and contact tracing on the epidemic dynamics in each state, we assumed a linear "ramp-up" of either testing and/or contact tracing from July 1th – 15th, 2020, after which both parameters remain constant. We then predicted the daily number of cases and deaths (Figures 3 and S10). We found that under current levels of reopening and control, at least 26 states would see a continuous increase in cases and deaths (Figure S10). Even with increased testing and contact tracing, some of these states will still experience a short-term increase in cases and deaths (Figures 3 and S10). For example, Ohio, Texas, and Washington may experience a substantial short-term increase of cases and deaths even if their current testing and contact tracing rate were doubled within the next two weeks (Figure 3B-D). Moreover, reported cases may slightly increase during the "ramp-up" period (Figure 3). We also found that in most states additional relaxation of restrictions without simultaneously increasing contact tracing may exacerbate disease dynamics and result in large-scale outbreaks (Figure S10).

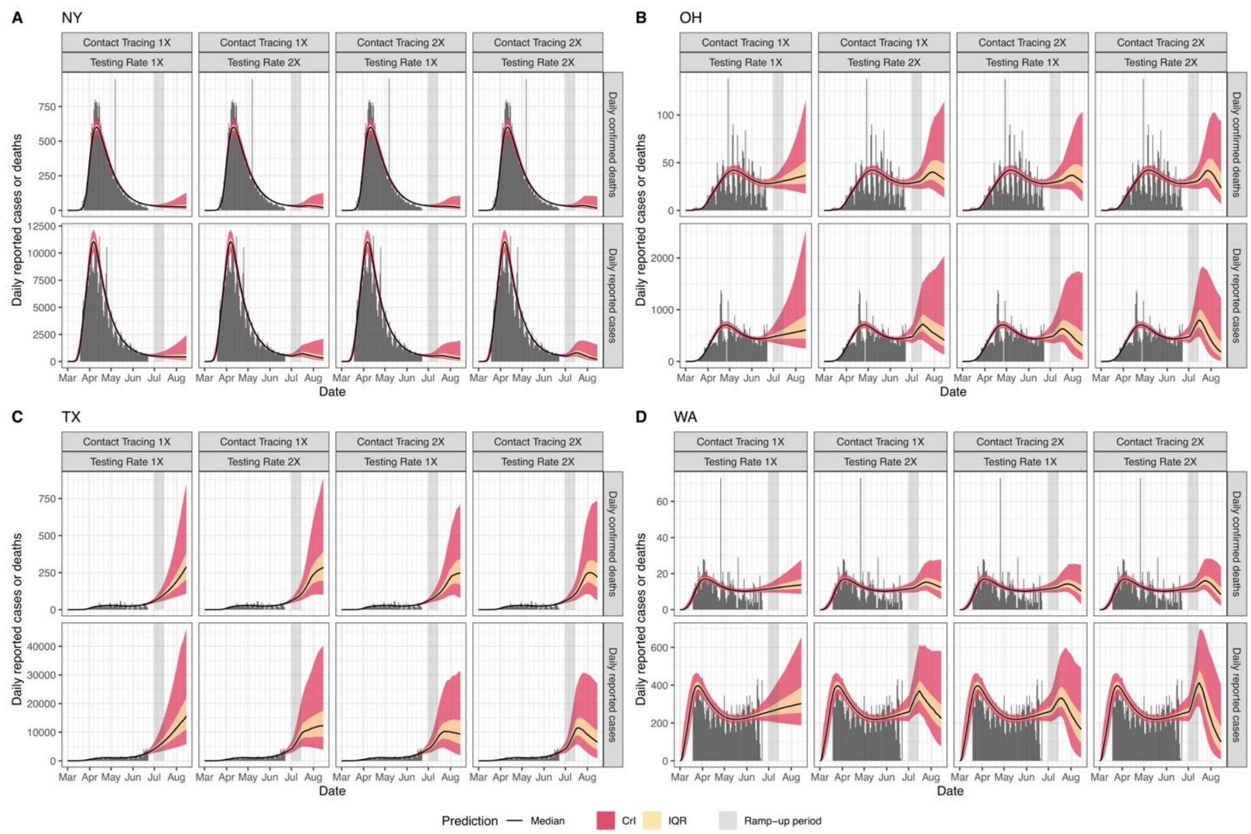

Figure 3. Predicted time-course (median, IQR, and 95% CrI) of daily reported cases and deaths under different testing and contact tracing rates (1X and 2X) in New York (A), Ohio (B), Texas (C), and Washington State (D).

We next evaluated the maximal degree of rebound in transmission (i.e., level of reopening) permitted while keeping R_{eff}<1 under different testing and contact tracing scenarios (Figure 4). We found that under the current level of testing and contact tracing rate, 36 states cannot keep their R_{eff}<1 even with only 25% reopening/rebound in transmission (Figure 4A). By doubling the current testing rate, four states (Connecticut, North Dakota, Nebraska, Rhode Island) could keep their R_{eff}<1 even with a 75% level of reopening (Figure 4B). By doubling contact tracing, five states (Connecticut, Delaware, Maryland, Pennsylvania, Rhode Island) could remove all mobility restrictions while keeping R_{eff}<1 (Figure 4C). By doubling both testing rate and contact tracing, 12 states could remove all mobility restrictions while keeping R_{eff}<1 (Figure 4D).

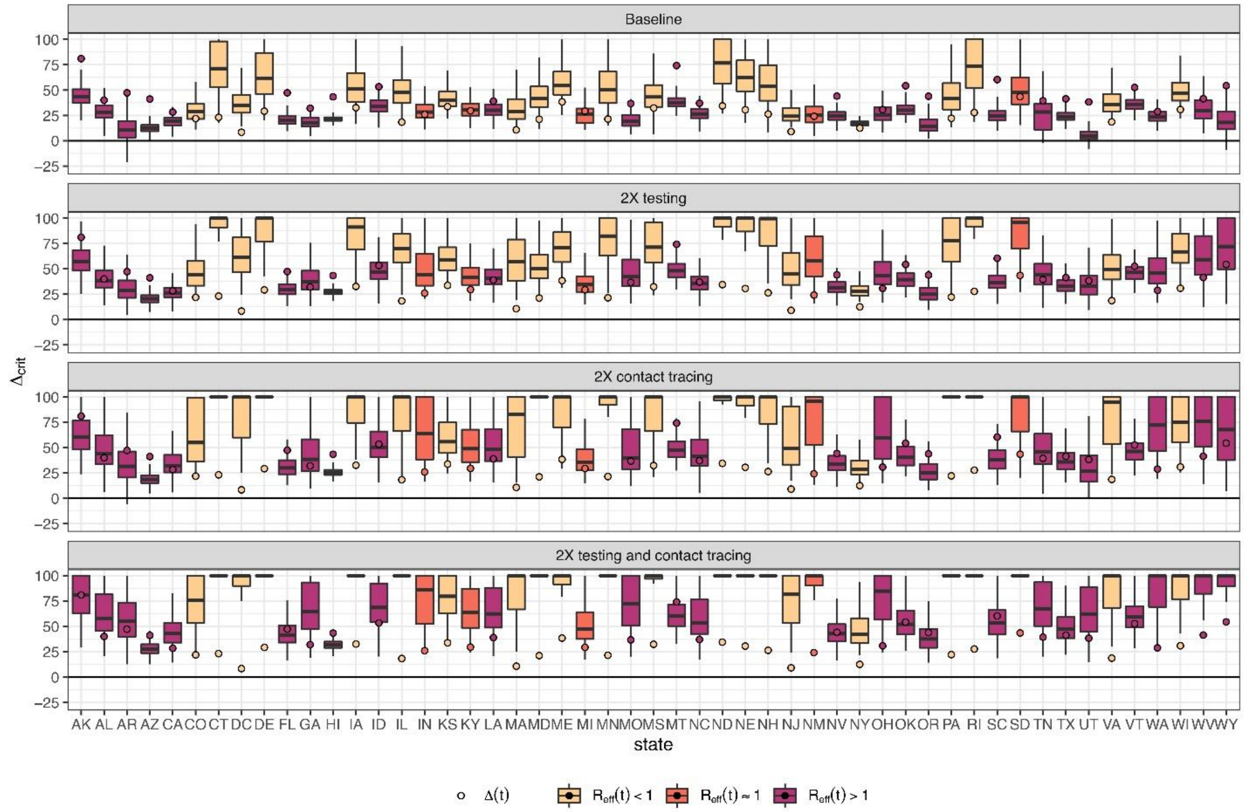

Figure 4. Reopening/rebound in transmission permitted (0 = minimum shelter-in-place value, 1 = return to no restrictions) to keep $R_{eff} < 1$ if (A) testing and contact rates are unchanged, (B) testing rate is doubled, (C) contact tracing is doubled, or (D) both testing and contact tracing are doubled. $\Delta(t)$ the level of reopening/rebound in transmission on June 20th, 2020 is shown by the circle. All boxplots show median, IQR, and 95% CrI.

We categorized states by the additional amount of mitigation efforts needed to keep $R_{eff} < 1$ with at least 75% confidence (Figures 5 and S10). We found that under current control efforts, three states (Connecticut, North Dakota, Rhode Island) could continue to curtail their epidemic curve even with an additional 25% reopening ("None" category), and that an additional 16 states and the District of Columbia could curtail their epidemic curve without additional reopening ("Low" category). 13 states could curtail their epidemic curve by doubling their current testing or contact tracing rate ("Moderate" category), while for 11 states by doubling both testing and contact tracing are need ("High" category). The remaining seven states (Alaska, Hawaii, Georgia, Florida, Oklahoma, Oregon, Montana) require not only doubling both testing and contact tracing, but also additional social distancing restrictions, in order to curtail their epidemic curve ("Very High" Category).

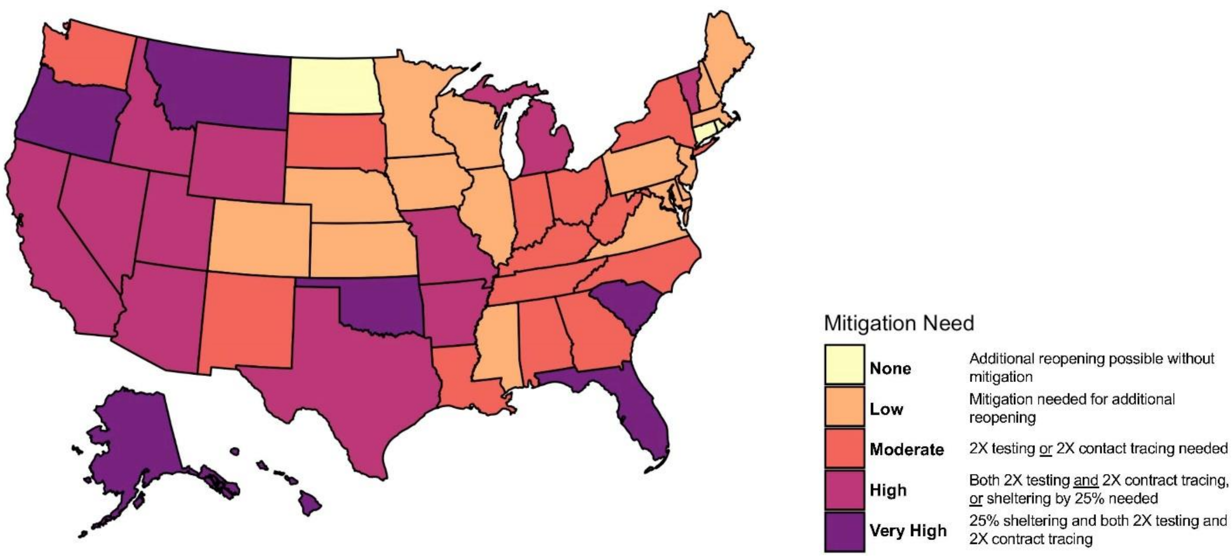

Figure 5. State-specific level of mitigation needed as a June 20, 2020 to curtail the spread of COVID-19 (keeping $R_{eff} < 1$ with at least 75% confidence, equivalent to the upper bound of the Interquartile range (IQR)).

Discussion

There is a delicate and continuous balance to strike between the use of social distancing measures to mitigate the spread of an emerging and deadly disease such as COVID-19 and the need for re/opening various sectors of activities for the social, economic, mental, and physical well-being of a community. To address this issue, it is imperative to design measurable, data-driven, and flexible milestones for identifying when to make specific transitions with regard to easing or retightening specific social distancing measures. We developed a data-driven SARS-CoV-2 transmission dynamic model not only to make short-term predictions on COVID-19 incidence and mortality in the US, but more importantly to evaluate the impact that relaxing social distancing measures and increasing testing and contact tracing would have on the epidemic in each state.

We showed that in most states, control strategies implemented during their “shelter-in-place” period were sufficient to contain the outbreak, defined as reducing and ultimately maintaining the effective reproductive number below 1 (R_{eff}<1). However, for the majority of states, our modelling suggests that “reopening” has proceeded too rapidly and/or without adequate testing and contact tracing to prevent a resurgence of the epidemic. Even in states with currently decreasing incidence and mortality, such as New York and New Jersey, additional relaxation of restrictions is likely to “bend the epidemic curve upwards.” However, our model predicts that a combination of increased testing, increased contact tracing, and/or scaling back reopening will be sufficient for curtailing the spread of COVID-19. Specifically, doubling of current testing and contact tracing rates would enable the vast majority of states to either maintain or increase the easing of social distancing restrictions in a “safe” manner. Increasing testing and contact tracing rates entails both increasing the number of tests performed per day as well as requiring early identification and isolation of COVID-19. This can be accomplished through active case detection via efficient contact tracing strategies. However, it should also be noted that increased testing and contact tracing will lead to a short-term increase in reported cases because a larger fraction of the infected population is being observed, and that several weeks may pass before these rates begin to show a decline. It is therefore imperative that policymakers and the public recognize that such a surge is actually a sign that testing and tracing efforts are succeeding, and to have the patience to wait several weeks before these successes are reflected as declining rates of reported cases.

Like all modeling studies, our study has several limitations due to modelling assumptions and the quality of available data. The initiation of social distancing measures, such as stay-at-home orders in the US, for mitigating the spread of COVID-19 has occurred concurrently with increased promotion and application of other NPIs such as hygiene practices (e.g. hand hygiene, surface cleaning, cough etiquette, and wearing of face mask). These hygiene practices coupled with the avoidance of physical contact whenever possible (keeping six feet apart) could impact the spread of COVID-19 by reducing both the risk of exposure and the risk of transmission of SARS-CoV-2 from infected patients [14,15]. Though our model explicitly accounts for the differential contribution of social distancing (mobility reduction) versus hygiene practices and physical distancing to reducing COVID-19 transmission, we assume that the impact of hygiene practices and physical distancing was a function of social distancing (mobility reduction). While cell phone mobility data may continue to accurately reflect the contact rates, the impact of enhanced hygiene practices is more difficult to measure independently. As several states are easing their social distancing requirements, especially their stay-at-home orders, compliance with hygiene practices would become even more important for reducing individuals' risk of getting or transmitting the pathogen. However, keeping a high population-level adherence to these measures is required to mitigate the spread of the COVID-19 epidemic in a city, state, or nation [16]. As states are reopening various aspects of their economy, data on compliance with enhance hygiene practices and physical distancing are needed to improve the estimation of these measures' population-level impact on reducing disease transmission.

Additionally, consistent with previous COVID-19 modeling studies [17–19], our model uses a simple functional form to model increases in testing rate from early March to June, 2020.

This testing rate was estimated through model fitting to daily reported case and mortality data. Particularly in states that have seen a substantial increase in testing capability and efforts during the month of May, our simple time varying assumption may underestimate the current level of testing and contact tracing. However, it should be noted that increased testing capacity does not necessarily lead to increased rate of testing if individuals are unaware, unwilling, or unable to be tested [20]. Having contact tracing and date of symptoms onset data would enable us to compute a better estimate the current testing and contact tracing rate in each state. Our also model assumes that all individuals who test positive to COVID-19 are effectively isolated for the rest for their infectious period and no longer contribute to disease transmission. Though voluntary compliance to COVID-19 self-quarantine recommendations may be high across the US, it is likely not 100%. Therefore, the assumption of effective isolation of all identified cases may cause our model to slightly overestimate the impact of increase testing rate on disease dynamics. However, we anticipate that this assumption would only have a marginal impact on the qualitative nature of our results. Finally, our model does not explicitly account for age-stratified risk of disease transmission and mortality. This age-stratification is important for designing and evaluating social distancing and testing strategies that are targeted towards the elderly population which are at higher risk of COVID-19-induced hospitalization and death [21]. As reopening the economy becomes an imperative for states across the US, age- or risk-targeted interventions may be a valuable tool to mitigate the burden of the pandemic. Future modeling studies could investigate the effectiveness of age- or risk-targeted non-pharmaceutical and potential pharmaceutical (vaccine or therapeutic) interventions for controlling the spread and burden of COVID-19.

In sum, we use a data-driven mathematical modeling approach to study the impacts of social distancing, testing, and contact tracing on the transmission dynamics of SARS-CoV-2. Our findings emphasize the importance for public health authorities not only to monitor the case and mortality dynamics of SARS-CoV-2 in their state, but also to understand the impact of their existing social distancing measures on SARS-CoV-2 transmission and evaluate the effectiveness of their testing and contact tracing programs for promptly identifying and isolating new cases of COVID-19. As reported case rates are increasing widely across US states because social distancing restrictions have been eased to allow more economic activity to resume, we find that most states need to either significantly scale back reopening or enhance their capacity and scale of testing, case isolation, and contact tracing programs in order to prevent large-scale increases in COVID-19 cases and deaths.

Online Methods

Our overall approach is as follows: 1) develop a mathematical model (an SEIR-type compartmental model) that incorporates social distancing data, case identification via testing, isolation of detected cases, and contact tracing; 2) assess the model's predictive performance by training (calibrating) it to reported cases and mortality data from March 19^{th} to April 30^{th}, 2020 and validating its predictions against data from May 1^{st} to June 20^{th}, 2020; and 3) use the model, trained on data through June 20^{th}, 2020, to predict future incidence and mortality. The final stage of our approach predicts future events under a set of scenarios that include increased case detection though expanded testing rate, contact tracing, and relaxation or increase of measures to promote social distancing. All model fitting is performed in a Bayesian framework in order to incorporate available prior information and address multivariate uncertainty in model parameters.

Model formulation

Our model is illustrated in Figure 1, with parameters and prior distributions listed in Table 1. We modified the standard SEIR model to address testing and contact tracing. In our model formulation I class also includes infectious pre-symptomatic individuals. With respect to testing, separate compartments were added for untested, "freely roaming" infected individuals (I_U), tested/isolated cases I_T, fatalities F_T. In balancing considerations of model fidelity and parameter identifiability, we made the reasonably conservative assumptions that all tested cases are effectively isolated (through self-quarantine or hospitalization) and thus unavailable for transmission, and that all COVID-related deaths are identified/tested.

With respect to contact tracing, the additional compartment S_C represents unexposed contacts, who undergo a period of isolation during which they are not susceptible before returning to S; while E_C and I_C represent contacts who were exposed. Again, the reasonably conservative assumption was made that all exposed contacts undergo testing, with an accelerated testing rate compared to the general population. We assume a closed population of constant size N for each state.

The ordinary differential equations governing our model are as follows:

$$\frac{dS}{dt} = -S \cdot c \cdot [\beta + (1-\beta) \cdot f_C] \cdot I_U/N + S_C \cdot \alpha$$

$$\frac{dS_C}{dt} = -S_C \cdot \alpha + S \cdot c \cdot (1-\beta) \cdot f_C \cdot I_U/N$$

$$\frac{dE}{dt} = -E \cdot \kappa + S \cdot c \cdot \beta \cdot (1 - f_C) \cdot I_U/N$$

$$\frac{dE_C}{dt} = -E_C \cdot \kappa + S \cdot c \cdot \beta \cdot f_C \cdot I_U/N$$

$$\frac{dI_U}{dt} = -I_U \cdot (\lambda + \rho) + E \cdot \kappa$$

$$\frac{dI_C}{dt} = -I_C \cdot (\lambda_C + \rho_C) + E_C \cdot \kappa$$

$$\frac{dR_U}{dt} = I_U \cdot \rho + I_C \cdot \rho_C$$

$$\frac{dI_T}{dt} = -I_T \cdot (\rho + \delta) + I_U \cdot \lambda + I_C \cdot \lambda_C$$

$$\frac{dR_T}{dt} = I_T \cdot \rho$$

$$\frac{dF_T}{dt} = I_T \cdot \delta$$

The testing rates λ and λ_C, the fatality rate δ, and the recovery rate of traced contacts ρ_C are each composites of several underlying parameters. The testing rate defined as

$$\lambda(t) = F_{test,0} \cdot \left[1 - \frac{1}{1 + e^{(t-T_T)/\tau_T}}\right] \cdot Sens_{test} \cdot k_{test},$$

where $F_{test,0}$ is the current testing coverage (fraction of infected individuals tested), $Sens_{test}$ is the test sensitivity (true positive rate), and k_{test} is rate of testing for those tested, with a typical time-to-test equal to $1/k_{test}$. The time-dependence term models the “ramp-up” of testing using a logistic function with a growth rate of $1/\tau_T$ days^{-1}, where T_T is the time where 50% of the current testing rate is achieved. Similarly, for testing of traced contacts, the same definition is used with the assumption that all identified contacts are tested, $F_{test,0} = 1$ and at a faster assumed testing rate $k_{C,test}$:

$$\lambda_C(t) = \left[1 - \frac{1}{1 + e^{(t-T_T)/\tau_T}}\right] \cdot Sens_{test} \cdot k_{C,test},$$

Because all contacts are assumed to be tested, the rate ρ_C at which they enter the "recovered" compartment R_U is simply the rate of false negative test results:

$$\rho_C(t) = \left[1 - \frac{1}{1 + e^{(t-T_T)/\tau_T}}\right] \cdot (1 - Sens_{test}) \cdot k_{test}$$

The fatality rate is adjusted to maintain consistency with the assumption that all COVID-19 deaths are identified, assuming a constant infected fatality rate (IFR). Specifically, we first calculated the fraction of infected that are tested and positive

$$f_{pos}(t) = f_C \frac{\lambda_C(t)}{\lambda_C(t) + \rho_C(t)} + (1 - f_C) \frac{\lambda(t)}{\lambda(t) + \rho}.$$

Where f_C is the fraction of contact identified through contact tracing.

Then the case fatality rate $CFR(t) = IFR/f_{pos}(t)$. Because the $CFR = \delta/(\delta + \rho)$, this implies

$$\delta(t) = \rho \frac{CFR(t)}{1 - CFR(t)} = \rho \frac{IFR}{f_{pos}(t) - IFR}.$$

The model is "seeded" $N_{initial}$ cases on February 29, 2020. Because in the early stages of the outbreak, there may be multiple "imported" cases, we only fit to data from March 19, 2020 onwards, one week after the U.S. travel ban was put in place [22].

Our model is fit to daily case y_c and death y_d data (cumulative data are not used for fitting because of autocorrelation). To adequately fit the case and mortality data, we accounted for two lag times. First, a lag is assumed between leaving the I_U compartment and public reporting of a positive test result, accounting for the time it takes to seek a test, obtaining testing, and have the result reported. No lag is assumed for tests from contact tracing. Second, a lag time is assumed between entering the fatally ill compartment F_T and publically reported deaths. Additionally, we use a negative binomial likelihood in order to account for the substantial day-to-day variation in reporting results. The corresponding equations are as follows:

$$y_{obs,[c,d]}(t) \sim NegBin[\alpha_{[c,d]}, p_{[c,d]}(t)]$$

$$p_{[c,d]}(t) = \frac{y_{pred,[c,d]}(t)}{\alpha_{[c,d]} + y_{pred,[c,d]}(t)}$$

$$y_{pred,c}(t) = I_U(t - \tau_{case}) \cdot \lambda(t) + I_C(t) \cdot \lambda_C(t)$$

$$y_{pred,d}(t) = I_T(t - \tau_{death}) \cdot \delta(t)$$

In this parameterization, as the shape parameter $\alpha \to \infty$, the likelihood becomes a Poisson distribution with expected value $y_{pred,[c,d]}$, whereas for small values of α there is substatial inter-individual variability. Case and death data were sourced from The COVID Tracking Project [23].

Finally, we derived time-dependent the time-dependent and effective reproduction numbers in this model, given by

$$R(t) = \frac{c \cdot \beta \cdot (1 - f_C)}{\lambda + \rho}$$

and

$$R_{eff}(t) = R(t) \cdot \frac{S(t)}{N}.$$

Incorporating social distancing, enhanced hygiene practices, and reopening

The impact of social distancing, hygiene practices, and reopening were modeled through a time-dependence in the contact rate c and the transmission probability per infected contact β:

$$c(t) = c_0 \cdot [\theta(t) + (1 - \theta_{min}) \cdot r(t)]$$

$$\beta(t) = \beta_0 \cdot \theta(t)^{\eta}$$

The $\theta(t)$ function parameterizes social distancing during the progression to shelter-in-place, and is modeled as a Weibull function

$$\theta(t) = \theta_{min} + (1 - \theta_{min}) e^{-(t/\tau_\theta)^{n_\theta}},$$

which starts a unity and decreases to θ_{min}, with T_θ being Weibull scale parameter and n_θ the Weibull shape parameter (Figure S1).

The $r(t)$ function parameterizes relative increase in contacts after shelter-in-place, with $r = 1$ corresponding to a return to baseline $c = c_0$.

$$r(t) = r_{max} \frac{t - \tau_\theta - \tau_s}{\tau_r} [u(t - t_r) - u(t - t_{rmax})] + u(t - t_{rmax})$$

$$u(t) = \text{Heaviside}(t) \approx 1 - \frac{1}{1 + e^{4t}}$$

$$t_r = \tau_\theta + \tau_s$$

$$t_{rmax} = \tau_\theta + \tau_s + \tau_r$$

The term $r(t)$ is 0 before t_r, linear between t_r and t_{rmax}, and constant at a value of r_{max} after that, and made continuous by approximating the Heaviside function by a logistic function. The reopening time is defined as τ_s days after τ_θ, and the maximum relative increase in contacts r_{max} happens τ_r days after that.

We selected the functional form above for $c(t)$ because it was found to be able to represent a wide variety of social distancing data, including cell phone mobility data from Unacast [24] and Google [25], as well as restaurant booking data from OpenTable [26] . We used these different mobility sources to derive state-specific prior distributions because different social distancing datasets had different values for θ_{min}, τ_θ, n_θ, τ_S, r_{max}, and τ_R (Figure S2).

With respect to the reduction in transmission probability β, we assumed that during the "shelter-in-place" phase, hygiene-based mitigation paralleled this decline with an effectiveness power η, and that this mitigation continued through re-opening.

Finally, we define an overall "reopening" parameter Δ that measures the "rebound" in disease transmission $c \cdot \beta$ relative to its minimum, defined to be 0 during shelter-in-place (i.e., $R(t)$ is at a minimum), and 1 when all restrictions are removed (when $R(t) = R_0$), which can be derived as:

$$\Delta(t) = \frac{c \cdot \beta/(c_0 \cdot \beta_0) - \theta_{min}^{1+\eta}}{1 - \theta_{min}^{1+\eta}}.$$

Scenario evaluation

We used the model to make several inferences about the current and future course of the pandemic in each state. First, we consider the effective reproduction number. Two time points of particular interest are the time of minimum R_{eff}, reflecting the degree to which shelter-in-place and other interventions were effective in reducing transmission, and the final time of the simulation, June 20, 2020, reflecting the extent to which reopening has increased R_{eff}. Additional parameters of interest are the current levels of reopening $\Delta(t)$, testing λ, and contact tracing f_C.

We then conducted scenario-based prospective predictions using our model's parameters as estimated through June 20, 2020. We asked the following questions:

(a) Assuming current levels of reopening, what increases in general testing λ and/or contact tracing f_C would be necessary to bring $R_{eff} < 1$?

(b) What amount reopening Δ can maintain $R_{eff} < 1$ under four different scenarios: current values of testing and contact tracing, doubling testing, double tracing, and doubling both testing and tracing?

(c) What will the rates of new cases and deaths be under different scenarios? Specifically, we evaluate the impact of increases in testing and contact tracing under current levels of reopening, as well as increases or decreases of 25%.

For (a), we evaluated the posterior probability that $R_{eff} < 1$ under scaling transformations $\lambda \to \lambda \cdot \mu_\lambda$ and $f_C \to f_C \cdot \mu_C$ with scaling factors μ_λ and μ_C:

$$R_{eff}(t) = \frac{S(t) \cdot c \cdot \beta \cdot (1 - \mu_C \cdot f_C)}{\mu_\lambda \cdot \lambda + \rho} = \frac{S(t) \cdot c_0 \cdot \beta_0 \cdot (1 - \mu_C \cdot f_C)}{\mu_\lambda \cdot \lambda + \rho} \left[\Delta \cdot (1 - \theta_{min}^{1+\eta}) + \theta_{min}^{1+\eta}\right]$$

For (b), we fixed the scaling factors at 1 or 2, and solved the above equation for Δ_{crit} such that $R_{eff} < 1$:

$$\Delta_{crit}(\mu_\lambda, \mu_C) = \left[\frac{\mu_\lambda \cdot \lambda + \rho}{S(t) \cdot c_0 \cdot \beta_0 \cdot (1 - \mu_C \cdot f_C)} - \theta_{min}^{1+\eta}\right] \frac{1}{1 - \theta_{min}^{1+\eta}}$$

Values of $\Delta_{crit} \geq \Delta(t)$ indicate the additional degree of reopening possible while maintaining $R_{eff} < 1$, while values of $\Delta_{crit} < \Delta(t)$ indicate a reduction of reopening is needed.

Finally, for (c), we additionally evaluated changes in reopening $\Delta \to \Delta + \Delta_\Delta$ for Δ_Δ values of +25% or -25%, for a total of 12 scenarios (4 different levels of testing and tracing, and 3 different levels of reopening). We then ran the SEIR model forward in time until August 31, 2020. For all three intervention parameters μ_C, μ_Λ, and Δ_Δ, we assumed a "ramp-up" period of 2 weeks from July 1-July 14, 2020. To summarize the relative urgency of mitigation in each state, we categorized states based on which scenarios resulted in the IQR of $R_{eff}(t)$ being < 1 on July 15, 2020.

Software and code:

Posterior distributions were sampled using Markov chain Monte Carlo simulation performed using MCSim version 6.1.0 using Metropolis within Gibbs sampling [27]. For each US state, four chains of 200,000 iterations each were run, with the first 20% of runs discarded, and 500 posterior samples saved for analysis. For each parameter, comparison of interchain and intrachain variability was assessed to determine convergence, with the potential scale reduction factor R ≤ 1.2 considered converged [28]. Additional analysis of model outputs was performed in RStudio version 1.2.1335 [29] with R version 3.6.1[30]. The codes used to generate our results will be available on Github prior to publication at https://github.com/wachiuphd/COVID-19-Bayesian-SEIR-US.

Data availability statement:

The following publicly available datasets are used:

- Mobility data from Unacast were sourced from https://covid19-scoreboard-api.unacastapis.com/api/search/covidstateaggregates_v3.
- Mobility data from Google were sourced from https://www.gstatic.com/covid19/mobility/Global_Mobility_Report.csv.
- Restaurant booking data were sourced from OpenTable https://www.opentable.com/state-of-industry.

- Case and death data were sourced from The COVID Tracking Project (https://covidtracking.com/).

Mobility data are shown in Supplemental Figure S2. Case and death data are shown in Figures 1 and 3, and Supplemental Figures S3-S6, S10. All data used are also available in the software and code repository.

Funding: National Science Foundation (NSF RAPID DEB 2028632) and National Institutes of Health, National Institute of Environmental Health Sciences (P30 ES029067).

Acknowledgements: We thank F.Y. Bois, J.K. Cetina, M. Giannoni, I. Rusyn, and W. Więcek for useful input and advice on scenario development, model formulation, and MCMC simulation. We also thank Unacast for making their mobility data available for researchers, and The COVID Tracking Project for compiling case and mortality data and providing it to the public. Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.

Table 1. Model inputs, parameters and prior distributions for Bayesian analysis.

Symbol	Definition (units)	Sampled parameter(s)	Prior [Truncation]	Notes
Pop	Population size	Input (not sampled)	Constant	[31]
N_{init}	Initial I_U on 2020-02-29	N_{init}	LogN(1000, 10) [1, 10000]	¶
$1/\alpha$	Self-isolation time after contact tracing	$T_{isolation} = 1/\alpha$	LogN(14, 2) [7, 21]	†
$1/\kappa$	Latent period (d)	$T_{latent} = 1/\kappa$	N(4,1) [2,7]	[32,33]
c_0	Baseline contact rate (contacts d^{-1})	c_0	N(13, 5) [7, 20]	[34]
ρ	Recovery rate (d^{-1})	$T_{recover}= 1/\rho$	LogN(10, 1.5) [5, 30]	[33,35]
β_0	Transmission rate (d^{-1})	$R_0 = c_0\beta_0/\rho$	N(2.9, 0.78) [1.46, 4.5]	[36–38]
f_C	Fraction of contacts traced (unitless)	f_C	LogN(0.25, 2) [0.15, 1]	[39]
T_T	Date of startup of testing (d)	T_T	N(70, 10) [60, 90]	¶
λ	General positive diagnosis rate (d^{-1})	$\lambda = F_{test}\, Sens_{test}\, k_{test}$	Derived	[36,40,41]
F_{test}	General test coverage (unitless)	F_{test}	N(0.5, 0.2) [0.2, 0.8]	[36,40,41]
$Sens_{test}$	Test sensitivity (unitless)	$Sens_{test}$	N(0.7, 0.1) [0.6, 0.95]	[42]
k_{test}	General testing rate (d^{-1})	$\tau_{test} = 1/k_{test}$	N(7, 3) [2, 12]	[43,44]
λ_C	Contacts positive diagnosis rate (d^{-1})	$\lambda_C = Sens_{test}\, k_{test,C}$	Derived	
$k_{C,test}$	Contacts testing rate (d^{-1})	$\tau_{C,test} = 1/k_{C,test}$	N(2, 1) [1, 3]	¶
ρ_C	Rate of infected contacts testing negative (d^{-1})	$\rho_C = (1 - Sens_{test})\, k_{test,C}$	Derived	
δ	Fatal illness rate (d^{-1})	IFR (infected fatality rate)*	LogN(0.01, 2) [0.001, 0.1]	[35,45]
θ_{min}	Minimum of $\theta(t)$	θ_{min}	Validation: Beta(2,2) Calibration: State-specific	¶ ɋ
τ_θ	Weibull scale parameter	τ_θ	Validation: N(21, 7) [7, 35] Calibration: State-specific	¶ ɋ
n_θ	Weibull shape parameter	n_θ	Validation: LogN(6, 2) [1,11] Calibration: State-specific	¶ ɋ
η	Hygiene effectiveness relative to social distancing (unitless)	η	Beta(2,2)	¶
τ_s	Duration of shelter in place (d)	τ_s	Validation: N(30, 30) [0, 90] Calibration: State-specific	[46]
τ_r	Duration of linear increase after shelter-in-place (d)	τ_r	Validation: N(45, 30) [0, 105] Calibration: State-specific	¶ ɋ
r_{max}	Maximum relative increase in contacts from shelter-in-place (unitless)	r_{max}	Validation: Beta(2,2) Calibration: State-specific	¶ ɋ

τ_{case}	Lag time for observing confirmed case	τ_{case}	LogN(7, 2) [1, 14]	¶
τ_{death}	Lag time for observing confirmed death	τ_{death}	LogN(7, 2) [1, 14]	¶
α_{pos}	Negative Binomial shape parameter for cases likelihood function	α_{pos}	LogU(4, 40)	¶
α_{death}	Negative Binomial shape parameter for deaths likelihood function	α_{death}	LogU(8, 40)	¶

LogN(GM, GSD) = lognormal distribution with geometric mean GM and geometric standard deviation GSD
N(M,SD) = normal distribution with mean M and standard deviation SD
U(MIN,MAX) = uniform distribution with minimum MIN and maximum MAX
LogU(MIN, MAX) = log-uniform distribution with minimum MIN and maximum MAX
Time (t) is measured from t=1 corresponds to 2020-01-01.
¶ Assumed, non-informative prior.
† Standard contact tracing guidance is to self-isolate for 2 weeks.
⁋ For calibration to 6/20/20, state-specific priors were derived by fitting to different social distancing data sets, with each parameter's mean, standard deviation, and range used to define a normal distribution prior.
* See Methods for relationship between IFR and δ.

References

1. Johns Hopkins Coronavirus Resource Center. Available at: https://coronavirus.jhu.edu/. (Accessed: 20th June 2020)

2. Holshue, M. L. *et al.* First Case of 2019 Novel Coronavirus in the United States. *N. Engl. J. Med.* **382**, 929–936 (2020).

3. Weiss, P. & Murdoch, D. R. Clinical course and mortality risk of severe COVID-19. *The Lancet* **395**, 1014–1015 (2020).

4. Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention. *JAMA - Journal of the American Medical Association* **323**, 1239–1242 (2020).

5. Kucharski, A. J. *et al.* Effectiveness of isolation, testing, contact tracing and physical distancing on reducing transmission of SARS-CoV-2 in different settings. *medRxiv* 2020.04.23.20077024 (2020). doi:10.1101/2020.04.23.20077024

6. Chu, D. K. *et al.* Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet* **0**, (2020).

7. What is social distancing and how can it slow the spread of COVID-19? | Hub. Available at: https://hub.jhu.edu/2020/03/13/what-is-social-distancing/. (Accessed: 20th June 2020)

8. Bialek, S. *et al.* Geographic Differences in COVID-19 Cases, Deaths, and Incidence — United States, February 12–April 7, 2020. *MMWR. Morb. Mortal. Wkly. Rep.* **69**, 465–471 (2020).

9. Siedner, M. J. *et al.* Social distancing to slow the U.S. COVID-19 epidemic: an interrupted time-series analysis. *medRxiv* 2020.04.03.20052373 (2020). doi:10.1101/2020.04.03.20052373

10. Wagner, A. B. *et al.* Social Distancing Has Merely Stabilized COVID-19 in the US. *medRxiv* 2020.04.27.20081836 (2020). doi:10.1101/2020.04.27.20081836

11. Lifting Social Distancing Measures in America: State Actions & Metrics | The Henry J. Kaiser Family Foundation. Available at: https://www.kff.org/coronavirus-policy-watch/lifting-social-distancing-measures-in-america-state-actions-metrics/. (Accessed: 20th June 2020)

12. Testing in the U.S. | CDC. Available at: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/testing-in-us.html. (Accessed: 20th June 2020)

13. Garnett, G. P., Cousens, S., Hallett, T. B., Steketee, R. & Walker, N. Mathematical models in the evaluation of health programmes. *The Lancet* **378**, 515–525 (2011).

14. Leung, N. H. L. *et al.* Respiratory virus shedding in exhaled breath and efficacy of face masks. *Nat. Med.* **26**, 676–680 (2020).

15. Greenhalgh, T., Schmid, M. B., Czypionka, T., Bassler, D. & Gruer, L. Face masks for the public during the covid-19 crisis. *BMJ* **369**, (2020).

16. Eikenberry, S. E. *et al.* To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infect. Dis. Model.* **5**, 293–308 (2020).

17. Pitzer, V. E. *et al.* The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. *medRxiv* 2020.04.20.20073338 (2020). doi:10.1101/2020.04.20.20073338

18. Yamana, T., Pei, S., Kandula, S. & Shaman, J. Projection of COVID-19 Cases and Deaths in the US as Individual States Re-open May 4,2020. *medRxiv* 2020.05.04.20090670 (2020). doi:10.1101/2020.05.04.20090670

19. Peirlinck, M., Linka, K., Sahli Costabal, F. & Kuhl, E. Outbreak dynamics of COVID-19 in China and the United States. *Biomech. Model. Mechanobiol.* 1–15 (2020). doi:10.1007/s10237-020-01332-5

20. U.S. states see COVID-19 testing supply improvements, but challenges abound - The Washington Post. Available at: https://www.washingtonpost.com/health/as-coronavirus-testing-expands-a-new-problem-arises-not-enough-people-to-test/2020/05/17/3f3297de-8bcd-11ea-8ac1-bfb250876b7a_story.html. (Accessed: 20th June 2020)

21. Keeling, M. J. *et al.* Predictions of COVID-19 dynamics in the UK: short-term forecasting and analysis of potential exit strategies. doi:10.1101/2020.05.10.20083683

22. Presidential Proclamation — Travel From Europe. Available at: https://travel.state.gov/content/travel/en/traveladvisories/presidential-proclamation--travel-from-europe.html. (Accessed: 20th June 2020)

23. The COVID Tracking Project | The COVID Tracking Project. Available at: https://covidtracking.com/. (Accessed: 20th June 2020)

24. Data for Good. Available at: https://www.unacast.com/data-for-good. (Accessed: 20th June 2020)

25. COVID-19 Community Mobility Reports. Available at: https://www.google.com/covid19/mobility/. (Accessed: 20th June 2020)

26. State of the Industry | OpenTable. Available at: https://www.opentable.com/state-of-industry. (Accessed: 20th June 2020)

27. Bois, F. Y. GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models. *Bioinformatics* **25**, 1453–1454 (2009).

28. Gelman, A. & Rubin, D. B. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* **7**, 457–472

29. RStudio | Open source & professional software for data science teams - RStudio. Available at: https://rstudio.com/. (Accessed: 20th June 2020)

30. R: The R Project for Statistical Computing. Available at: https://www.r-project.org/.

31. programs-surveys/popest/datasets/2010-2019/state/detail. Available at: https://www2.census.gov/programs-surveys/popest/datasets/2010-2019/state/detail/. (Accessed: 20th June 2020)

32. Lauer, S. A. *et al.* The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* (2020). doi:10.7326/M20-0504

33. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).

34. Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).

35. Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet. Infect. Dis.* **0**, (2020).

36. Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).

37. Riou, J. & Althaus, C. L. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Eurosurveillance* **25**, 2000058 (2020).

38. Park, S. W. *et al.* Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: a new framework and applications to the novel coronavirus (2019-nCoV) outbreak. *medRxiv* 2020.01.30.20019877 (2020). doi:10.1101/2020.01.30.20019877

39. Covid Act Now. Available at: https://covidactnow.org/?s=49762. (Accessed: 20th June 2020)

40. Arons, M. M. *et al.* Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. *N. Engl. J. Med.* (2020). doi:10.1056/nejmoa2008457

41. Nishiura, H. *et al.* Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). (2020). doi:10.1016/j.ijid.2020.03.020

42. Watson, J., Whiting, P. F. & Brush, J. E. Interpreting a covid-19 test result. *BMJ* **369**, m1808 (2020).

43. Cummings, M. J. *et al.* Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *Lancet* **395**,

1763–1770 (2020).

44. Sun, K., Chen, J. & Viboud, C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *Lancet Digit. Heal.* **2**, e201–e208 (2020).

45. Onder, G., Rezza, G. & Brusaferro, S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA - Journal of the American Medical Association* **323**, 1775–1776 (2020).

46. Map: How states are reopening after coronavirus shutdown - Washington Post. Available at: https://www.washingtonpost.com/graphics/2020/national/states-reopening-coronavirus-map/. (Accessed: 20th June 2020)