# The challenges of modeling and forecasting the spread of COVID-19

Andrea L. Bertozzi[a,b,1] (ORCID), Elisa Franco[b,c], George Mohler[d], Martin B. Short[e], and Daniel Sledge[f]

[a]Department of Mathematics, University of California, Los Angeles, CA 90095; [b]Department of Mechanical and Aerospace Engineering, University of California, Los Angeles, CA 90095; [c]Department of Bioengineering, University of California, Los Angeles, CA 90095; [d]Department of Computer Science, Indiana University–Purdue University Indianapolis, Indianapolis, IN 46202; [e]Department of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332; and [f]Department of Political Science, University of Texas at Arlington, Arlington, TX 76019

The coronavirus disease 2019 (COVID-19) pandemic has placed epidemic modeling at the forefront of worldwide public policy making. Nonetheless, modeling and forecasting the spread of COVID-19 remains a challenge. Here, we detail three regional-scale models for forecasting and assessing the course of the pandemic. This work demonstrates the utility of parsimonious models for early-time data and provides an accessible framework for generating policy-relevant insights into its course. We show how these models can be connected to each other and to time series data for a particular region. Capable of measuring and forecasting the impacts of social distancing, these models highlight the dangers of relaxing nonpharmaceutical public health interventions in the absence of a vaccine or antiviral therapies.

COVID-19 | pandemic | branching process | compartmental models

The world is in the midst of an ongoing pandemic, caused by the emergence of a novel coronavirus. Pharmaceutical interventions such as vaccination and antiviral drugs are not currently available. Over the next year, addressing the coronavirus disease 2019 (COVID-19) outbreak will depend critically on the successful implementation of public health measures including social distancing, shelter in place orders, disease surveillance, contact tracing, isolation, and quarantine (1, 2). On 16 March, Imperial College London released a report (3) predicting dire consequences if the United States and the United Kingdom did not swiftly take action against the pandemic. In both nations, governments responded by implementing more stringent social distancing regulations (4). We now have substantially more case data from the United States, as well as the benefit of analyses performed by scientists and researchers across the world (5–12). Nonetheless, modeling and forecasting the spread of COVID-19 remain a challenge.

Here, we present three basic models of disease transmission that can be fit to data emerging from local and national governments. While the Imperial College study employed an agent-based method (one that simulates individuals getting sick and recovering through contacts with other individuals in the population), we present three macroscopic models: 1) exponential growth, 2) self-exciting branching process, and 3) the susceptible–infected–resistant (SIR) compartment model. These models have been chosen for their simplicity, minimal number of parameters, and for their ability to describe regional-scale aspects of the pandemic. In presenting these models, we demonstrate how they are connected and note that in different cases one model may fit better than another. Because these models are parsimonious, they are particularly well suited to isolating key features of the pandemic and to developing policy-relevant insights. We order them according to their usefulness at different stages of the pandemic—exponential growth for the initial stage, a self-exciting branching process when one is still analyzing individual count data going into the development of the pandemic, and a macroscopic mean-field model going into the peak of the

disease. The branching process can also track changes over time in the dynamic reproductive number.

These models highlight the significance of fully implemented and sustained social distancing measures. Put in place at an early stage, distancing measures that reduce the virus's reproduction number—the expected number of individuals who an infected person will spread the disease to—may allow much-needed time for the development of pharmaceutical interventions. By slowing the speed of transmission, such measures may also reduce the strain on health care systems and allow for higher-quality treatment for those who become infected. Importantly, the economic consequences of such measures may lead political leaders to consider relaxing them. The models presented here, however, demonstrate that relaxing these measures in the absence of pharmaceutical interventions may allow the pandemic to reemerge. Where this takes place, social distancing efforts that appear to have succeeded in the short term will have little impact on the total number of infections expected over the course of the pandemic.

The epidemiological perspective on modeling infectious disease spread involves consideration of a larger number of

> **Significance**
>
> The coronavirus disease 2019 (COVID-19) pandemic has placed epidemic modeling at the forefront of worldwide public policy making. Nonetheless, modeling and forecasting the spread of COVID-19 remain a challenge. Here, we present and detail three regional-scale models for forecasting and assessing the course of the pandemic. This work is intended to demonstrate the utility of parsimonious models for understanding the pandemic and to provide an accessible framework for generating policy-relevant insights into its course. We show how these models can be connected to each other and to time series data for a particular region. Capable of measuring and forecasting the impacts of social distancing, these models highlight the dangers of relaxing nonpharmaceutical public health interventions in the absence of a vaccine or antiviral therapies.

modeling parameters detailing the spread of and recovery from the disease, additional compartments corresponding to age categories, and other related choices (e.g., refs. 3 and 13). A data-driven approach to modeling COVID-19 has also emerged, in which statistical and machine learning models are used for forecasting cases, hospitalizations, deaths, and impacts of social distancing (14, 15). Our work demonstrates the utility of parsimonious epidemic models for understanding the pandemic and provides an accessible framework for a larger group of quantitative scientists to follow and forecast the COVID-19 pandemic. It includes explanations that will help allow scientific researchers to develop insights that may contribute to public health policy making, including contributing to public health forecasting teams. Importantly, the branching process model that we detail is relatively new and underutilized in epidemiology. It provides a method for quantitatively estimating dynamic reproduction numbers, which can be critical in assessing the impact of distancing measures over time. The results from the parsimonious models presented here are consistent with recent analyses from public health officials in California (16) and with the original Imperial College model (3).

We present examples of forecasts for viral transmission in the United States. Where other studies have typically developed and presented one model (choosing to fit parameters within the chosen model), our analysis compares three different forecasting models using a fitting criterion. The results of these models differ depending on whether the data employed cover confirmed cases or mortality. In addition, many aspects of disease spread, such as incubation periods, fraction of asymptomatic but contagious individuals, seasonal effects, and the time between severe illness and death, are not considered here. In some cases (e.g., seasonal), relevant data do not exist, and in other cases (e.g., age of patients), we choose not to include additional parameters in favor of parsimony.

## Results

**A. Exponential Growth.** Epidemics naturally exhibit exponential behavior in the early stages of an outbreak, when the number of infections is already substantial but recoveries and deaths are still negligible. If at a given time $t$ there are $I(t)$ infected individuals and $\alpha$ is the rate constant at which they infect others, then at early times (neglecting recovered individuals), $I(t) = I_0 e^{\alpha t}$. The time it takes to double the number of cumulative infections (doubling time) is a common measure of how fast the contagion spreads: if we start with $\bar{I}$ infections, then at time $T_d = \ln 2/\alpha$ we achieve $2\bar{I}$ infections. For the COVID-19 outbreak, exponential growth is seen in data from multiple countries (Fig. 1), with remarkably similar doubling times in the early stages of the epidemic. For COVID-19, we expect an exponential growth phase during the first 15 to 20 d of the outbreak in the absence of public health interventions such as social distancing, isolation, or quarantine. This estimate is based on patient data from the Wuhan outbreak, which indicate that the average time from illness onset to death or discharge is between 17 and 21 d for hospitalized patients (20, 21). Because deaths are a fraction of infections, they initially increase at a similar exponential pace,
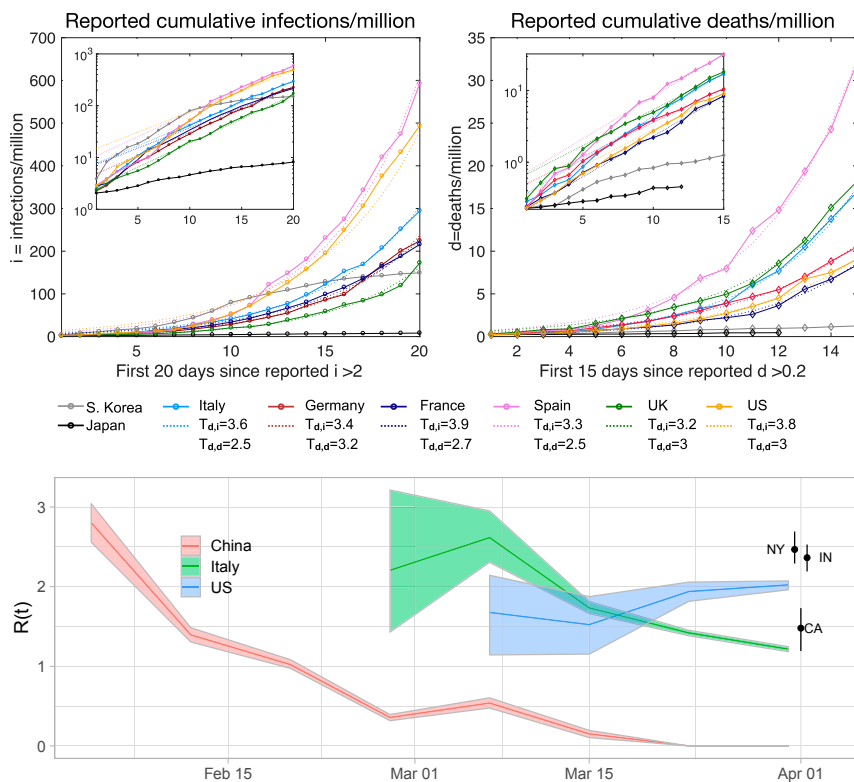


**Fig. 1.** (*Upper*) Exponential model applied to new infection and death data for Italy, Germany, France, Spain, the United Kingdom, and the United States, normalized by the total country population (54). *Insets* show the same data on a logarithmic scale. Both the normalized infection $i$ and death $d$ data were thresholded to comparable initial conditions for each country; fits are to the first 15 to 20 d of the epidemic after exceeding the threshold. The fitted doubling time is shown for both infections ($T_{d,i}$) and death ($T_{d,d}$) data. Data from Japan and South Korea are shown for comparison. (*Lower*) Dynamic reproduction number (mean and 95% CI) of COVID-19 for China, Italy, and the United States estimated from reported deaths (17) using a nonparametric branching process (18). Current estimates are as of 1 April 2020 of the reproduction number in New York (NY), California (CA), and Indiana (IN; confirmed cases used instead of mortality for Indiana). Reproduction numbers of COVID-19 vary in different studies and regions of the world (in addition to over time) but have generally been found to be between 1.5 and 6 (19) prior to social distancing.

with some delay relative to the beginning of the outbreak. These observed doubling time estimates (2 to 4 d) are significantly smaller than early estimates ($\sim$7 d) obtained using data collected in Wuhan (22).

**B. Self-Exciting Branching Process.** A branching point process (23–25) can also model the rate of infections over time. Point process models are data driven and allow for parametric or nonparametric estimation of the reproduction number and transmission timescale. They also facilitate estimation of the probability of extinction at early stages of an epidemic. These models have been used for various social interactions including spread of Ebola (26), retaliatory gang crimes (27), and email traffic (28, 29). The intensity (rate) of infections can be modeled as

$$\lambda(t) = \mu + \sum_{t_i < t} \mathcal{R}(t_i) w(t - t_i), \qquad \text{[1]}$$

where $t$ is the current time and $t_i$ are the times of previous infection incidents. Here, the dimensionless reproduction number, $\mathcal{R}(t)$, evolves in time (18, 30–33) to reflect changes in disease reproduction in response to public health interventions (e.g., school closings, social distancing, closures of nonessential businesses, isolation, and quarantine). The distribution of interevent times $w(t_i - t_j)$ is a gamma or Weibull distribution (11, 33, 34) with shape parameter $k$ and scale parameter $b$. Finally, the parameter $\mu$ allows for exogenous infection cases. The point process in Eq. **1** is an approximation to the common SIR model of infectious diseases (described later) during the initial phase of an epidemic when the total number of infections is small compared with the overall population size (35).

Given Eq. **1**, the quantity

$$p_{ij} = \mathcal{R}(t_j) w(t_i - t_j) / \lambda(t_i) \qquad \text{[2]}$$

gives the probability of secondary infection $i$ having been caused by primary infection $j$. The dynamic reproduction number $\mathcal{R}(t)$ can then be estimated via expectation–maximization (18) using a histogram estimator:

$$\mathcal{R}(t) = \sum_{k=1}^{B} r_k 1\{t \in I_k\}. \qquad \text{[3]}$$

Here, $I_k$ are intervals discretizing time, and $B$ is the number of such intervals. The reproduction number, $r_k$, in each interval $k$ is determined by

$$r_k = \sum_{t_i > t_j} p_{ij} 1\{t_j \in I_k\} / N_k, \qquad \text{[4]}$$

where $N_k$ is the total number of events in interval $k$.

Fig. 1, *Lower* shows the estimated dynamic reproduction number (36, 37) of COVID-19 in China, Italy, and the United States during the early stage of the pandemic, from late January 2020 to early April 2020. The branching point process is fit to mortality data (17) using an expectation–maximization algorithm (18). Public health measures undertaken in China appear to have reduced $\mathcal{R}(t)$ to below the self-sustaining level of $\mathcal{R} = 1$ by the middle of February. In Italy, public health measures brought the local value of $\mathcal{R}(t)$ down; as of early April, however, it remained above $\mathcal{R} = 1$. The estimated reproduction number in the United States as a whole stood at around 2.5. The reproduction number, however, varies notably by location.

This model can be adapted to capture the long-term evolution of the pandemic by incorporating a prefactor that accounts for the dynamic decrease in the number of susceptible individuals (35):

$$\lambda^h(t) = (1 - I_c(t)/N)(\mu + \sum_{t_i < t} \mathcal{R}(t_i) w(t - t_i)). \qquad \text{[5]}$$

Here, $I_c(t)$ is the cumulative number of infections that have occurred up to time $t$, and $N$ is the total population size. This version of the branching process model, referred to as HawkesN, represents a stochastic version of the SIR model; with large $\mathcal{R}$, the results of HawkesN are essentially deterministic. When projecting, we use our estimated $\mathcal{R}(t_i)$ at the last known point for all times going forward. Since the $N_t$ term is the number of infections, if our estimates for $\mathcal{R}(t_i)$ are based on mortality numbers, we must also choose a mortality rate to interpolate between the two counts; although existing estimated rates vary significantly, we choose 1% as a plausible baseline (38) (discussion in *Materials and Methods*). Alternatively, we also create forecasts for three US states based on fits to reported case data (Table 1).

**C. Compartmental Models.** The SIR model (40–42) describes a classic "compartmental" model with SIR population groups. A related model, susceptible–exposed–infected–resistant (SEIR), includes an "exposed" compartment that models a delay between exposure and infectiousness. The SEIR model was shown to fit historical death record data from the 1918 influenza epidemic (43), during which governments implemented extensive social distancing measures, including bans on public events, school closures, and quarantine and isolation measures. The SIR model can be fit to the predictions made in ref. 3 for agent-based simulations of the United States. The SIR model assumes a population of size $N$ where $S$ is the total number of susceptible individuals, $I$ is the number of infected individuals, and $R$ is the number of resistant individuals. For simplicity of modeling, we view deaths as a subset of resistant individuals, and deaths can be estimated from the dynamics of $R$; this is reasonable for a disease with a relatively small death rate. We also assume a timescale short enough such that humans' natural resistance to the disease does not introduce new susceptible people after recovery.

The SIR model equations are

$$\frac{dS}{dt} = -\beta \frac{IS}{N}, \qquad \frac{dI}{dt} = \beta \frac{IS}{N} - \gamma I, \qquad \frac{dR}{dt} = \gamma I. \qquad \text{[6]}$$

Here, $\beta$ is the transmission rate constant, $\gamma$ is the recovery rate constant, and $\mathcal{R}_0 = \beta/\gamma$ is the reproduction number. One integrates Eq. **6** forward in time from initial values of $S$, $I$, and $R$ at time 0. The SEIR model includes an exposed category $E$:

$$\frac{dS}{dt} = -\beta \frac{IS}{N}, \qquad \frac{dE}{dt} = \beta \frac{IS}{N} - aE,$$

$$\frac{dI}{dt} = aE - \gamma I, \qquad \frac{dR}{dt} = \gamma I.$$

Here, $a$ is the inverse of the average incubation time. Both models are fit, using maximum likelihood estimation with a Poisson likelihood, to data for three US states (California, New York, and Indiana) (17). Table 1 compares the results with the branching process. We use the relative likelihood based upon the Akaike Information Criterion (AIC) (39) to measure model performance for each dataset; AIC is biased against models with more parameters. The SEIR model performs better on the confirmed data for California and Indiana, possibly due to the larger amount of data, compared with mortality for which SIR is the best for all three states. The branching process performs best for confirmed cases in New York. Our choice of fitting follows the method in ref. 43 for the 1918 pandemic death data. Our focus is on model comparison rather

## Table 1. Model Fits to Early Time Data

| | Data | Model | $\mathcal{R}_0$ | $\gamma$ | $I_0$ | $a$ | Peak | logL | Rel. L |
|---|---|---|---|---|---|---|---|---|---|
| CA | conf. | SIR | 2.4 | 0.14 | 0.025 | – | 5/16/20 | 43230 | 0.14 |
| CA | conf. | SEIR | 4.9 | 0.10 | 0.05 | 0.28 | 5/16/20 | 43233 | 1.00 |
| CA | conf. | Branching | 1.6 | $(b,k,\mu)=(2.73,4.0,1.60)$ | | | 5/21/20 | 42722 | 0.00 |
| CA | mort. | SIR | 2.7 | 0.12 | 0.10 | – | 5/5/20 | 252 | 1.00 |
| CA | mort. | SEIR | 3.1 | 0.20 | 0.10 | 0.38 | 5/3/20 | 252 | 0.37 |
| CA | mort. | Branching | 1.5 | $(b,k,\mu)=(1.73,3.31,.04)$ | | | 5/8/20 | 242 | 0.00 |
| IN | conf. | SIR | 4.4 | 0.06 | 0.005 | – | 5/13/20 | 9098 | 0.00 |
| IN | conf. | SEIR | 4.7 | 0.16 | 0.005 | 0.20 | 5/9/20 | 9116 | 1.00 |
| IN | conf. | Branching | 2.4 | $(b,k,\mu)=(3.18,4.0,0.17)$ | | | 5/1/20 | 8947 | 0.00 |
| IN | mort. | SIR | 3.7 | 0.09 | 0.005 | – | 4/26/20 | 19.95 | 1.00 |
| IN | mort. | SEIR | 4.6 | 0.14 | 0.005 | 0.36 | 4/25/20 | 19.96 | 0.37 |
| IN | mort. | Branching | 2.1 | $(b,k,\mu)=(3.26,2.45,.02)$ | | | 5/1/20 | 10.68 | 0.00 |
| NY | conf. | SIR | 2.1 | 0.19 | 0.10 | – | 4/27/20 | 569616 | 0.00 |
| NY | conf. | SEIR | 3.4 | 0.19 | 0.10 | 0.36 | 4/25/20 | 569452 | 0.00 |
| NY | conf. | Branching | 1.5 | $(b,k,\mu)=(1.90,4.0,3.69)$ | | | 4/28/20 | 571203 | 1.00 |
| NY | mort. | SIR | 4.1 | 0.10 | 0.005 | – | 4/12/20 | 6313 | 1.00 |
| NY | mort. | SEIR | 5.0 | 0.20 | 0.005 | 0.32 | 4/12/20 | 6313 | 0.37 |
| NY | mort. | Branching | 2.5 | $(b,k,\mu)=(3.25,3.97,.04)$ | | | 4/18/20 | 6272 | 0.00 |



(*Left*) Fit of confirmed case (conf.) or mortality (mort.) data from California (CA), Indiana (IN), and New York (NY) states to three different models (SIR, SEIR, and branching process) using Poisson regression. The table shows the log likelihood and the relative likelihood $\exp((AIC_{min} - AIC) = 2)$ based on the Akaike information criteria (AIC) (39). The blue lettering corresponds to the lowest AIC value. The branching process parameters include a Weibull shape $k$ and scale $b$ for $w(t)$, along with the exogenous rate $\mu$. The table also shows parameters from the fit and the projected date for the peak in new cases for each of these datasets; the projected peak date for the branching process is made using the HawkesN model. For each state, we run the fit on both confirmed case data and mortality data, taken from ref. 17. (*Right*) Shown are the actual data points compared with the fitted curves.
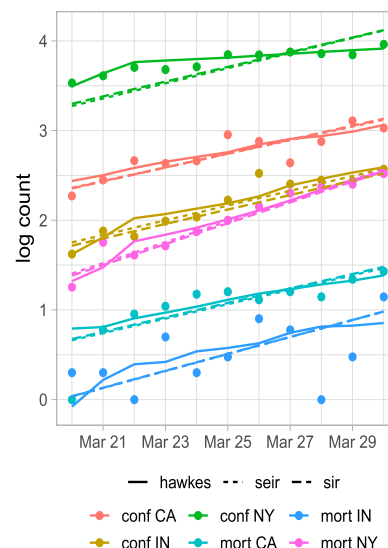
than measuring uncertainty of parameters in a specific model, as is currently being done for hospital demand forecasts in Los Angeles (16). For interval forecasts with uncertainty quantification, one may consider a negative binomial alternative to Poisson regression that captures overdispersion in case and death counts (44, 45).

We can further understand the role of parameters in our models via a dimensionless formulation of Eq. **6**. There are two timescales dictated by $\beta$ and $\gamma$. Therefore, if time is rescaled by $\gamma$ to $\tau = \gamma t$ and $s = S/N$, $i = I/N$, and $r = R/N$ represent fractions of the population in each compartment, then in the case of a novel outbreak with no initially resistant individuals, we obtain

$$\frac{ds}{d\tau} = -\mathcal{R}_0 is, \quad \frac{di}{d\tau} = \mathcal{R}_0 is - i, \quad \frac{dr}{d\tau} = i,$$
$$(s, i, r)|_{\tau=0} = (1 - \epsilon, \epsilon, 0), \qquad \textbf{[7]}$$

where $0 < \epsilon << 1$ is the initial fraction of the infected population at the start time, and the system retains only one dimensionless parameter $\mathcal{R}_0$ that, in conjunction with the initial conditions, completely determines the resulting behavior. For Eq. **7**, the shapes of the solution curves $s(\tau), i(\tau), r(\tau)$ do not depend on $\epsilon$, other than exhibiting a time shift that depends logarithmically on $\epsilon$ (Fig. 2). This is a universal solution for the SIR model in the limit of small $\epsilon$ (Fig. 3), depending only on $\mathcal{R}_0$. Critically, the height of the peak in $i(t)$ and the total number of resistant/susceptible people by the end of the epidemic are determined by $\mathcal{R}_0$ alone. However, the sensitivity of the time translation to the parameter $\epsilon$ and the dependence of true time values of the peak on parameter $\gamma$ make SIR challenging to fit to data at the early stages of an epidemic when Poisson statistics and missing information are prevalent. When using early-time death data to fit SIR, the estimate of the percentage of deaths per total number of infections (chosen as 1% here) has a sensitivity that can be understood directly in terms of this shift in the time to peak. This is important information for public health officials,

policy makers, and for political leaders interested in decreasing $\mathcal{R}_0$ for substantial periods of time. This sensitivity to parameters helps explain why projections of the outbreak can display such large variability and highlights the need for extensive disease testing within the population to more accurately track the epidemic curve. After the surge in infections, the model asymptotes to an end state in which $r$ approaches the end value $r_\infty$ and $s$ approaches $1 - r_\infty$ and the infected population approaches 0. The value $r_\infty$ satisfies a well-known transcendental equation (46–48). A phase diagram of the universal solutions for several $\mathcal{R}_0$ values is shown in Fig. 2, *Upper Right*. The dynamics start in the bottom right corner where $s$ is almost 1 and follow the colored line to terminate on the $i = 0$ axis at the value $s_\infty$. A rigorous derivation of the limiting state under the assumptions here can be found in refs. 46–48.

Under the SIR (and similar) model(s), if $\mathcal{R}_0$ is decreased during the middle of an outbreak, through social distancing and other public health measures, the rate of new infections will decrease. However, unless the number of infected individuals is brought down to zero, the outbreak will likely reemerge, and the total number of infections may still be a large fraction of the population. This is illustrated in Fig. 3, where we present scenarios of no social distancing vs. short-term social distancing with parameters from Table 1 for death data up through the end of March, fit to the SIR model. We caution that the goal of these scenarios is not to produce highly accurate percentages but rather, to present scenarios under different basic assumptions that illustrate the usefulness of social distancing measures and the potential danger in easing them too soon.

For the New York state scenario, with $\mathcal{R}_0$ presumed to decrease by a factor of two with distancing measures, the outbreak is not completely controlled by distancing, and the number of infections continues to rise to approximately 10% by mid-April. In contrast, without distancing measures the scenario shows four times that number of infected individuals by mid-April, a level that would have represented a potentially disastrous
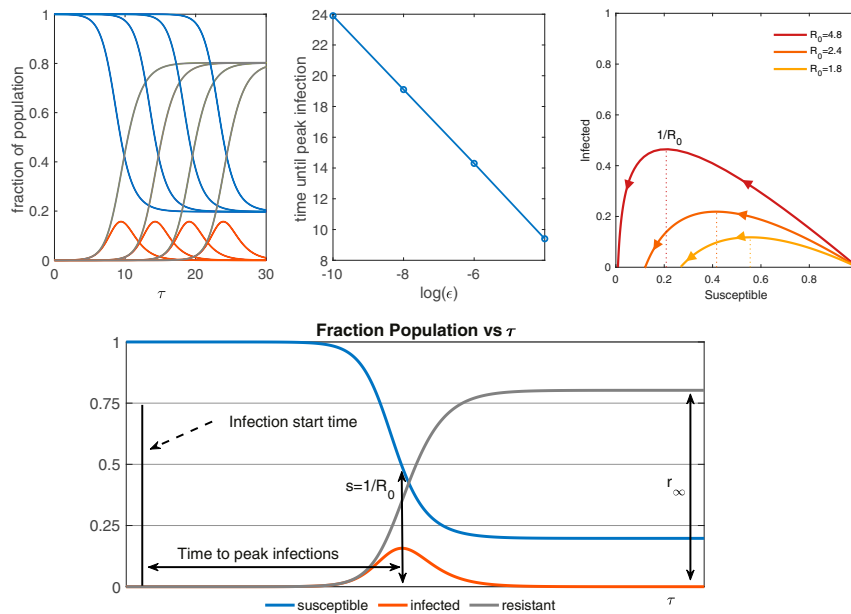
**Fig. 2.** Solution of the dimensionless SIR model (5) with $\mathcal{R}_0 = 2$. *Upper Left* shows the graphs of $s$ (blue), $i$ (orange), and $r$ (gray) on the vertical axis vs. $\tau$ on the horizontal axis, for different $\epsilon$. The corresponding values of $\epsilon$ from left to right are $10^{-4}$, $10^{-6}$, $10^{-8}$, $10^{-10}$, respectively. *Upper Center* shows the time until peak infections vs. $\log(\epsilon)$ for the values shown in *Upper Left*. This asymptotic tail to the left makes it challenging to fit data to SIR in the early stages. *Upper Right* is a phase diagram for fraction of infected vs. fraction of susceptible with the direction of increasing $\tau$ indicated by arrows, for three different values of $\mathcal{R}_0$. *Lower* displays a typical set of SIR solution curves over the course of an epidemic, with important quantities labeled.

strain on the hospital system. In the California scenario, distancing measures bring the effective $\mathcal{R}_0$ closer to one, thus keeping infections at a much smaller portion of the population than the scenario with no social distancing, at least during the period when the distancing is still in effect. We take these scenarios one step further by suddenly stopping distancing on 5 May (this is both extreme and hypothetical, but it serves to illustrate the model). Because of the low initial infected count in California, bringing $\mathcal{R}_0$ back to the original predistancing level produces a curve that follows the original peak scenario, just shifted later in time. For New York state, because a nontrivial fraction of the population is initially infected, there are fewer to infect, and the new peak is less steep than the scenario without any distancing.

## Discussion

Our analysis, employing parsimonious models, illustrates several key points. 1) The reproduction number $\mathcal{R}$ is highly variable both over time and by location, and this variability is compounded by distancing measures. These variations can be calculated using a stochastic model, and lower $\mathcal{R}$ is critical to decreasing strains on health care systems and to creating time to develop effective vaccines and antiviral therapies. 2) Mor-

tality data and confirmed case data have statistics that vary by location and by time depending on testing and on accurate accounting of deaths due to the disease. Differences in collection methods and in the accuracy of morbidity and mortality data can lead to different projected outcomes. 3) Nonpharmaceutical public health interventions (NPIs) such as social distancing and shelter in place orders offer an important means of reducing the virus's reproduction number. Nonetheless, NPIs may not have a substantial impact on the total number of infections unless sustained over time. Policy makers should be cautious about scaling back distancing measures after early signs of effectiveness.

During the 1918 influenza pandemic, the early relaxation of social distancing measures led to a swift uptick in deaths in some US cities (43). The models presented here help to explain why this effect might occur, as illustrated in Fig. 3. Already, policy makers in many jurisdictions have started to implement new social protocols that allow for increased economic activity. In the United States, where public health authority is vested largely in states and localities, key decisions about such measures will be in the hands of local officials, with national agencies such as the Centers for Disease Control and Prevention playing a coordinating role and offering guidance (49). Nationally, policy
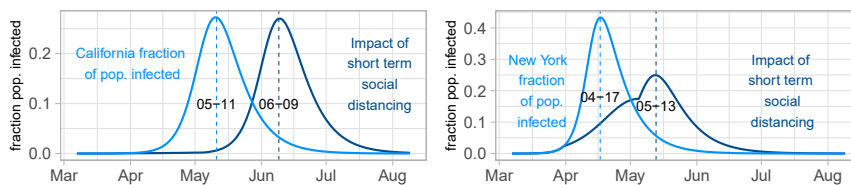


**Fig. 3.** Scenarios for the impact of short-term social distancing: fraction of population vs. date. (*Left*) California SIR model based on mortality data with parameters from Table 1 ($\mathcal{R}_0 = 2.7$, $\gamma = .12$, $I_0 = .1$) under two scenarios: $\mathcal{R}_0$ constant in time (light blue) and $\mathcal{R}_0$ cut in half from 27 March (1 wk from the start of the California shutdown) to 5 May but then returned to its original value, to represent a short-term distancing strategy (dark blue). (*Right*) New York SIR model with parameters from Table 1 ($\mathcal{R}_0 = 4.1$, $\gamma = .1$, $I_0 = 05$) under the same two scenarios but with short-term distancing occurring over the dates of 30 March (1 wk from the start of the New York shutdown) to 5 May. In both states, the distancing measures suppress the curve and push the peak infected date into the future, but the total number of cases is only slightly reduced.

makers may also consider funding mechanisms and regulatory measures that might facilitate a more unified approach to the pandemic (50), as well as fiscal measures aimed at ameliorating the economic effects of social distancing and shelter in place orders.

Of the models presented here, only the point process model tracks the reproductive number changing over time (although compartmental models can be modified to do the same). Dynamic change in the reproductive number is the major challenge in forecasting such a pandemic. Some new approaches to forecasting long-term trends in the COVID-19 pandemic attempt to address this concern by fitting curves in countries in the later stages of spread (after social distancing) and then applying those fitted models to regions in earlier stages (15). The long-term SIR curves in Fig. 3 are idealized counterfactuals in the event of no social distancing, rather than forecasts of the actual peak date and total infected. Moreover, in the absence of distancing policies, people may still choose to distance out of fear driven by an upswing in recent deaths, as was modeled in ref. 43 in the 1918 pandemic.

The models presented here are parsimonious, making a variety of assumptions in order to increase understanding and to avoid overfitting the limited and incomplete data available; more complex models have been introduced and are currently in use (3, 13). Even with these simple models, however, the parameters obtained from our fits (Table 1) can vary significantly for a given location. Although we have in each case determined which of these fits appears to have most validity, in many cases these are not strong indicators. These models have several sources of uncertainty, including parameter uncertainty, variation based on data or model type used, and most importantly, uncertainty in the severity and length of social distancing measures, which can change the peak date by months or even create multiple peaks. This variability in outcomes highlights the challenges of modeling and forecasting the course of a pandemic during its early stages and with only limited data. This uncertainty is a major challenge for policy makers, who must consider the social and economic consequences of disruptive public health interventions while recognizing that relaxing them may swiftly lead to the reemergence of a devastating disease.

## Materials and Methods

**Relation between the Exponential Model and Compartment Models.** The exponential model is appropriate during the first stages of the outbreak, when recoveries and deaths are negligible: in this case, the SIR compartment model can be directly reduced to an exponential model. If we assume $S \approx N$ in Eq. **6**, then $dI(t)/dt \approx (\beta - \gamma)I$, with the exponential solution $I(t) = I_0 e^{\alpha t}$ with $\alpha = \beta - \gamma$ and $I_0$ the initial number of infections. We expect at very early times $t \ll 1/\gamma$ that the recovery will lag infections so one might see $\alpha \sim \beta$ at very early times and then reduce to $\alpha \sim \beta - \gamma$ after $t > 1/\gamma$. Reports and graphs disseminated by the media typically report cumulative infections, which include recoveries and deaths. Using the SIR model, the total cumulative infections are $I_c(t) = I(t) + R(t)$ and evolve as $dI_c(t)/dt = \beta sI$. Integrating this, we see that $I_c$ likewise grows exponentially with the same rate $\alpha = \beta - \gamma$. An important observation is that the doubling time for cumulative infections $[T_d = \ln(2)/\alpha]$ will change during the early times, with a shorter doubling time when $t \ll 1/\gamma$ and a longer doubling time when $t > 1/\gamma$.

**Relation between the HawkesN and SIR Models.** Following refs. 35 and 51, first a stochastic SIR model can be defined where a counting process $I_c(t) = N - S(t)$ tracks the total number of infections up to time $t$, $N$ is the population size, and $S(t)$ is the number of susceptible individuals. The process satisfies

$$P(dC(t) = 1) = \beta S(t)I(t)dt/N + o(dt)$$
$$P(dR(t) = 1) = \gamma I(t)dt + o(dt),$$

which then gives the rate of new infections and new recoveries as (35)

$$\lambda^I(t) = \beta S(t)I(t)/N, \quad \lambda^R(t) = \gamma I(t).$$

It is shown in ref. 51 that the continuum limit of the counting process approaches the solution to the SIR model in Eq. **6**. Furthermore, for an exponential kernel $w(t)$ in the HawkesN model in Eq. **5** with parameter $\gamma$ and constant reproduction number ($\mathcal{R}_0$), then $E[\lambda^I(t)] = \lambda^H(t)$ where $\mu = 0$, $\beta = \mathcal{R}_0 \gamma$ (ref. 35 has further details).

**Fitting the SIR, SEIR, and Branching Process Models.** The parameters for SIR and SEIR in Table 1 were found using maximum Poisson likelihood regression (as in ref. 43 for death data from the 1918 pandemic in US cities) via grid search with ranges $I_0 \in [.005, .1]$, $\mathcal{R}_0 \in [1.5, 5]$, $\gamma \in [.01, .2]$, and $\mu \in [.01, .4]$.

The branching process was fit using a nonparametric expectation maximization algorithm, the details of which can be found in ref. 18. Models were fit to empirical new infections per day or new mortality counts per day. We assumed that 1% of those labeled as resistant in the simulations were fatal cases. For morality rates in the range of 0.3 to 3%, the peak date changes by up to 2 wk. Data quality is an issue during the COVID-19 pandemic due to lack of uniform testing. Uncertainty in estimates of the total number infected leads to uncertainty in forecasting results. The likelihood is somewhat flat at the maximum for models in Table 1, with multiple parameter combinations yielding plausible fits. Fitting SIR type models is known to be challenging due to parameter identifiability issues (52, 53). However, the peak date only varies by 1 to 3 d for parameters within two units of the maximum log likelihood.

## Data Sources

The data used in this manuscript were downloaded on 1 April 2020 from ref. 54 for Fig. 1, *Upper*, 16 June 2020 from ref. 17 for Fig. 1, *Lower*, and on 2 April 2020 from ref. 17 for Table 1. Data from both refs. 17 and 54 are publicly available. Case and death counts for COVID-19, such as those reported in ref. 17 that are used in the present study, are well known to suffer from ascertainment issues (55). The data also have other sources of error due to the nature in which they are aggregated from primary sources, where reporting is often lagged. Policy decisions based upon models fit to these data must take these ascertainment and data quality issues into account. The code and data used to generate Fig. 1, *Upper* can be downloaded from GitHub at https://github.com/francoelisa/PNAS2020. The code to download data and generate Fig. 1, *Lower* can be downloaded from GitHub at https://github.com/gomohler/pnas2020/tree/master/dynamic_R. The code to generate Table 1 is available in GitHub at https://github.com/gomohler/pnas2020.

1. A. L. Phelan, R. Katz, L. O. Gostin, The novel coronavirus originating in Wuhan, China: Challenges for global health governance. *JAMA* **323**, 709–710 (2020).
2. A. Pan et al., Association of public health interventions with the Epidemiology of the COVID-19 outbreak in Wuhan, China. *JAMA* **323**, 1–9 (2020).
3. N. M. Ferguson et al., "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand" (Report 9, Imperial College London, London, United Kingdom, 2020; https://doi.org/10.25561/77482).
4. M. Landler, S. Castle, Behind the virus report that jarred the U.S. and the U.K. to action. *NY Times*, 17 March 2020. https://www.nytimes.com/2020/03/17/world/europe/coronavirus-imperial-college-johnson.html. Accessed 24 March 2020.
5. N. Imai et al., "Transmissibility of 2019-nCoV" (Report 3, Imperial College London, London, United Kingdom, 2020; https://doi.org/10.25561/77148).
6. T. Li, Simulating the spread of epidemics in China on the multi-layer transportation network: Beyond the coronavirus in Wuhan. arXiv:2002.12280 (27 February 2020).
7. J. Riou, C. L. Althaus, Pattern of early human-to-human transmission of Wuhan 2019-NCOV. https://doi.org/10.1101/2020.01.23.917351 (24 January 2020).
8. T. A. Perkins et al., Estimating unobserved SARS-CoV-2 infections in the United States. https://doi.org/10.1101/2020.03.15.20036582 (26 March 2020).
9. A. J. Kucharski et al., Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *Lancet Infect. Dis.* **20**, 553–558 (2020).

10. L. C. Tindale *et al.*, Transmission interval estimates suggest pre-symptomatic spread of COVID-19. https://doi.org/10.1101/2020.03.03.20029983 (6 March 2020).

11. J. Hellewell *et al.*, Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Global Health* **8**, 488–496 (2020).

12. J. T. Wu, K. Leung, P. G. M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study *Lancet* **395**, 689–697 (2020).

13. A. Arenas *et al.*, A mathematical model for the spatiotemporal epidemic spreading of COVID19. https://doi.org/10.1101/2020.03.21.20040022 (23 March 2020).

14. N. Altieri *et al.*, Curating a COVID-19 data repository and forecasting county-level death counts in the United States. arXiv:2005.07882 (16 May 2020).

15. IHME COVID-19 Health Service Utilization Forecasting Team, C. J. L. Murray, Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. https://www.medrxiv.org/content/early/2020/03/30/2020.03.27.20043752 (30 March 2020).

16. R. J. Lewis *et al.*, Projections of hospital based healthcare demand due to COVID-19 in Los Angeles county. County DHS COVID 19 Predictive Modeling Team: May 6 update. http://dhs.lacounty.gov/wps/portal/dhs/covid19/mp. Accessed 10 April 2020.

17. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *Lancet* **20**, 533–534 (2020).

18. G. Mohler, F. Schoenberg, M. B. Short, D. Sledge, Analyzing the world-wide impact of public health interventions on the transmission dynamics of COVID-19. arXiv:2004.01714 (3 April 2020).

19. Y. Liu, A. A. Gayle, A. Wilder-Smith, J. Rocklöv, The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J. Trav. Med.* **27**, taaa021 (2020).

20. F. Pan *et al.*, Time course of lung changes on chest CT during recovery from coronavirus disease 2019 (COVID-19). *Radiology*, **295**, 715–721 (2020).

21. F. Zhou *et al.*, Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet* **395**, 1054–1062 (2020).

22. Q. Li *et al.*, Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).

23. S. Meyer, L. Held, M. Höhle, Spatio-temporal analysis of epidemic phenomena using the R package surveillance. *J. Stat. Softw.* **77**, 6981 (2017).

24. C. Farrington, M. Kanaan, N. Gay, Branching process models for surveillance of infectious diseases controlled by mass vaccination. *Biostatistics* **4**, 279–295 (2003).

25. F. Schoenberg, M. Hoffmann, R. Harrigan, A recursive point process model for infectious diseases. *Ann. Inst. Stat. Math.* **71**, 1271–1287 (2019).

26. R. Harrigan *et al.*, Real-time predictions of the 2018-2019 Ebola virus disease outbreak in the Democratic Republic of Congo using Hawkes point process models. *Epidemics* **28**, 100354 (2019).

27. A. Stomakhin, M. B. Short, A. L. Bertozzi, Reconstruction of missing data in social networks based on temporal patterns of interactions. *Inverse Probl.* **27**, 115013 (2011).

28. E. W. Fox *et al.*, Modeling e-mail networks and inferring leadership using self-exciting point processes. *J. Am. Stat. Assoc.* **111**, 564–584 (2016).

29. J. R. Zipkin, F. P. Schoenberg, K. Coronges, A. L. Bertozzi, Point-process models of social network interactions: Parameter estimation and missing data recovery. *Eur. J. Appl. Math.* **27**, 502–529 (2016).

30. S. Cauchemez *et al.*, Real-time estimates in early detection of SARS. *Emerg. Infect. Dis.* **12**, 110–113 (2006).

31. J. Wallinga, P. Teunis, Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am. J. Epidemiol.* **160**, 509–516 (2004).

32. L. Forsberg White, M. Pagano, A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Stat. Med.* **27**, 2999–3016 (2008).

33. T. Obadia, R. Haneef, P. Y. Boëlle, The $R_0$ package: A toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC Med. Inform. Decis. Mak.* **12**, 147 (2012).

34. B. J. Cowling *et al.*, The effective reproduction number of pandemic influenza: Prospective estimation. *Epidemiology* **21**, 842–846 (2010).

35. M. A. Rizoiu, S. Mishra, Q. Kong, M. Carman, L. Xie, SIR-Hawkes: Linking epidemic models and Hawkes processes to model diffusions in finite populations. https://arxiv.org/abs/1711.01679 (21 February 2018).

36. C. You *et al.*, Estimation of the time-varying reproduction number of COVID-19 outbreak in China. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3539694 (19 February 2020).

37. S. Riley *et al.*, Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science* **300**, 1961–1966 (2003).

38. R. Verity *et al.*, Estimates of the severity of COVID-19 disease. https://doi.org/10.1101/2020.03.09.20033357 (13 March 2020).

39. H. Akaike, A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974).

40. F. Brauer, C. Castillo-Chavez, Z. Feng, *Mathematical Models in Epidemiology* (Springer, 2019).

41. W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. A* **115**, 700–721 (1927).

42. J. Tolles, T. Luong, Modeling epidemics with compartmental models. *JAMA*, 10.1001/jama.2020.8420 (2020).

43. M. C. J. Bootsma, N. M. Ferguson, The effect of public health measures on the 1918 influenza pandemic in U.S. cities. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7588–7593 (2007).

44. G. Frasso, P. Lambert, Bayesian inference in an extended SEIR model with nonparametric disease transmission rate: An application to the Ebola epidemic in Sierra Leone. *Biostatistics* **17**, 779–792 (2016).

45. A. G. Angelov, M. Slavtchova-Bojkova, Bayesian estimation of the offspring mean in branching processes: Application to infectious disease data. *Comput. Math. Appl.* **64**, 229–235 (2012).

46. J. Miller, A note on the derivation of epidemic final sizes. *Bull. Math. Biol.* **74**, 2125–2141 (2012).

47. Miller J., Mathematical models of SIR disease spread with combined non-sexual and sexual transmission routes. *Infect. Dis. Model.* **2**, 35–55 (2017).

48. T. Harko, F. S. N. Lobo, M. K. Mak, Exact analytical solutions of the susceptible-infected-recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Appl. Math. Comput.* **236**, 184–194 (2014).

49. D. Sledge, *Health Divided: Public Health and Individual Medicine in the Making of the Modern American State* (University Press of Kansas, Lawrence, KS, 2017).

50. R. L. Haffajee, M. M. Mello, Thinking globally, acting locally—the U.S. response to COVID-19. *N. Engl. J. Med.* **382**, e75 (2020).

51. P. Yan, "Distribution theory, stochastic processes and infectious disease modelling" in *Mathematical Epidemiology*, F. Brauer, P. van den Driessche, J. Wu, Eds. (Lecture Notes in Mathematics, vol. 1945, Springer, Berlin, 2008), pp. 229–293.

52. N. D. Evans, L. J. White, M. J. Chapman, K. R. Godfrey, M. J. Chappell, The structural identifiability of the susceptible infected recovered model with seasonal forcing. *Math. Biosci.* **194**, 175–197 (2005).

53. K. Roosa, G. Chowell, Assessing parameter identifiability in compartmental dynamic models using a computational approach: Application to infectious disease transmission models. *Theor. Biol. Med. Model.* **16**, 1 (2019).

54. World Health Organization, Coronavirus disease (COVID-2019) situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. Accessed 1 April 2020.

55. J. T. Wu *et al.*, Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat. Med.* **26**, 506–510 (2020).